

Thu Dec 16 12:41:39 1999

us38183-1-38183.pep

Page 1

US3A183-1-38183
MIDIDYKEFGVATLISFLSDPEFSPBRLDLSAALYREALSEPHNSPHNLRALALICMELMTLAT
WGNVLEPASEADRLVYSVYVNTNGLFROLLMFHSCILPGRVYEVLSFGWJITPAIPYAPPAJLT
SLPTEPTVYRRGSPRRPRTSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRR
SLMDIDYKEFGVATLISFLSDPEFSPBRLDLSAALYREALSEPHNSPHNLRALALICMELMT
LATVWGNLEPASEADRLVYSVYVNTNGLFROLLMFHSCILPGRVYEVLSFGWJITPAIPYAPPA
PLISLPEPTVYRRGSPRRPRTSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRRSPRR

 N O T E S

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_bp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 12:43:38 1999; Maspar time 17.14 Seconds

Tabular output not generated. 491.237 Million cell updates/sec

Title: >US38183-1-38183
 Description: (1-396) from us38183-1-38183.pep
 Perfect Score: 2977
 Sequence: 1 MDIDPYKEFGATVELLSFLP.....RRRRSQSPRRRSQSRSSOC 396
 Scoring table: PAM 150
 Gap 11

Searched: 170751 seqs; 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.376; Variance 170.298; scale 0.196

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	46.2	194	32	W50242	Hepatitis B virus pre
2	1375	46.2	212	32	W50250	Hepatitis B virus pre
3	1375	46.2	346	5	R27473	S12/core protein.
4	1375	46.2	397	20	W09048	Plasmid pHEV DN Aa en
5	1358	45.6	184	1	P80959	Hepatitis B virus sub
6	1358	45.6	183	32	W50251	Hepatitis B virus p21
7	1358	45.6	193	32	W50241	Hepatitis B virus pre
8	1353	45.4	183	20	W09044	Hepatitis B virus cor
9	1349	45.3	183	1	R05635	Hepatitis B antigen.
10	1349	45.3	183	5	P00041	Sequence of core anti
11	1338	44.9	186	5	P00004	Sequence of core anti
12	1336	44.9	196	8	R40808	Hepatitis B core / PV
13	1337	44.9	208	8	R40808	Hepatitis B core / PV
14	1334	44.8	183	8	P40311	Hepatitis B virus core
15	1334	44.8	183	8	R40805	Hepatitis B core prot
16	1334	44.8	183	13	R68868	Hepatitis B virus pol

ID	W50242: standard; Protein; 194 AA.	ALIGNMENTS
17	1334	44.8
18	1333	44.8
19	1338	44.6
20	1335	44.5
21	1334	44.5
22	1334	44.5
23	1334	44.5
24	1334	44.5
25	1335	44.5
26	1335	44.5
27	1331	44.4
28	1330	44.3
29	1337	44.2
30	1317	44.2
31	1314	44.1
32	1313	44.1
33	1311	44.0
34	1311	44.0
35	1311	44.0
36	1309	44.0
37	1308	43.9
38	1308	43.9
39	1308	43.9
40	1303	43.8
41	1302	43.7
42	1287	43.2
43	1263	42.4
44	1259	42.3
45	1258	42.3

RESULT 1
 ID W50242: standard; Protein; 194 AA.
 AC W50242:1998 (first entry)
 DT 28-SEP-1998
 DE Hepatitis B virus precore p22 polypeptide Met-p22.
 KW Hepatitis B virus precore; inhibitor; HBV; nucleocapsid; gene therapy;
 KW hepatocytic; liver; Met-p22.
 OS Hepatitis B virus.
 FH Synthetic.
 FT Key
 FT Protein
 FT 2.194
 FT /label- p22
 PN W09809649-A1.
 PD 12-MAR-1998.
 PF 03-SEP-1997; U15500.
 PR (GHEO) GEN HOSPITAL CORP.
 PA Melegari M, Scaglioni PP, Wands JR;
 PI WPI: 98-193325/17.
 DR DNA encoding proteins which can be incorporated with wild type
 nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 11: Page 40; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein with an added N-terminal Met residue. p22 is produced by
 CC elimination of the 19-amino acid leader peptide from the 25 kDa
 CC full-length HBV precore protein (see W50250). Evidence is provided
 CC that HBV replication is inhibited in the presence of high levels of
 CC HBV precore or precore-related proteins. These proteins can be
 CC incorporated into HBV nucleocapsids along with the p21 core protein
 CC (see W50251), which is the usual nucleocapsid component, and
 CC thereby render the nucleocapsids deficient in encapsidating HBV
 CC pregenomic RNA. Thus, over-expression of the precore proteins, or
 CC certain variants of them, leads to transdominant inhibition of HBV
 CC replication. Suitable inhibitory proteins include p25 (see W50250),
 CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
 CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously

CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SO Sequence 194 AA;

Query Match 46.2%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 1.53e-101;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmdidykefatvellsfipstffpsvrdltdtasalyrealsepncphthalrgal 70
 :|||||
 QY 213 SMDIDPKEGATVELLSFIPSDFFPSVRDLDTASALYRALESPEHCSPHHTALRGAI 272
 |||||
 Db 71 lwcgelmrlatwgvnledpasrdlvsyvnltmgikfrqlwfhiscitfgretvieyl 130
 :|||||
 QY 273 lwcgelmrlatwgvnledpasrdlvsyvnltmgikfrqlwfhiscitfgretvieyl 332
 :|||||
 Db 131 vsfgvwlrtppayrppnapllstlpettvrrrgsprrrrrrrrrsgsprrrrrsgsr 190
 :|||||
 QY 333 vsfgvwlrtppayrppnapllstlpettvrrrgsprrrrrrrrrsgsprrrrrsgsr 392
 :|||||
 Db 191 esgc 194
 :|||||
 QY 393 ESQC 396

RESULT 2
 ID W50250 standard; Protein; 212 AA.
 AC W50250;

DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p25 polypeptide.
 KW Hepatitis B virus; inhibitor; HBV; nucleocapsid; gene therapy;
 KM Hepatocyte; liver; p25 protein.
 OS Hepatitis B virus.
 PN W09809649-A1.
 PD 12-MAR-1998.
 PE 03-SEP-1997; 015500.
 PR 03-SEP-1996; US-025370.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PT Claim 15, Page 35; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SO Sequence 212 AA;

Query Match 46.2%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 1.53e-101;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmdidykefatvellsfipstffpsvrdltdtasalyrealsepncphthalrgal 88
 :|||||
 QY 213 SMDIDPKEGATVELLSFIPSDFFPSVRDLDTASALYRALESPEHCSPHHTALRGAI 272
 :|||||

Db 89 lwcgelmrlatwgvnledpasrdlvsyvnltmgikfrqlwfhiscitfgretvieyl 148
 :|||||
 QY 273 lwcgelmrlatwgvnledpasrdlvsyvnltmgikfrqlwfhiscitfgretvieyl 332
 :|||||
 Db 149 vsfgvwlrtppayrppnapllstlpettvrrrgsprrrrrrrrrsgsprrrrrsgsr 208
 :|||||
 QY 333 vsfgvwlrtppayrppnapllstlpettvrrrgsprrrrrrrrrsgsprrrrrsgsr 392
 :|||||
 Db 209 esgc 212
 :|||||
 QY 393 ESQC 396

RESULT 3
 ID R27473 standard; Protein; 346 AA.
 AC R27473;

DT 24-FEB-1993 (first entry)
 DE S12/core protein.
 KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KW vaccinia virus; 13L; promoter; NYVAC; recombinant; HBV L;
 KW large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
 KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KW deletion loci; recipient loci.
 OS Synthetic.

FT Key Location/Qualifiers
 FT region 1..108
 FT /label S1
 FT region 109..163
 FT /label S2
 FT region 164..346
 FT /label Core

PN W09215672-A.
 PD 17-SEP-1992.
 PE 09-MAR-1992; 001906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRRO-) VIROGENETICS CORP.
 PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Lambach KJ, Norton EK, Paoletti E, Perkins ME, Pincus SE,
 PI Riviere M, Tarantola J, Taylor J;
 DR WPI: 92-331718/40.
 DR N-PDB: Q29105.

PT Vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure: Fig 13: 456pp; English.
 CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the 13L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NYVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, lpsAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NYVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC 035501-864.

Query Match 46.2%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 1.53e-101;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 nmdidykefatvellsfipstffpsvrdltdtasalyrealsepncphthalrgal 222

03-SEP-1997; U15500.
03-SEP-1996; US-025370.
(GCHO) GEN HOSPITAL CORP.
Melegari M, Scaglioni PP, Wands JR;
Wp1; 98-193325/17.
DNA encoding proteins which can be incorporated with wild type
nucleocapsid subunit(s) into a viral nucleocapsid - useful for
PT inhibition of viral replication, especially hepatitis B virus
PS Disclosure: Page 41-42: 60pp: English.
CC This is the 21 kDa (p21) protein of hepatitis B virus (HBV) that
CC assembled into a 180 Ksa subunit nucleocapsid structure that
CC promotes viral replication. Evidence is provided that HBV
CC replication is inhibited in the presence of high levels of HBV
CC precore or precore-related proteins. These proteins can be
CC incorporated into HBV nucleocapsids along with the p21 core protein
CC and thereby render the nucleocapsids deficient in encapsidating HBV
CC progenomic RNA. Thus, over-expression of the precore proteins or
CC certain variants of them, leads to transdominant inhibition of HBV
CC replication. Suitable inhibitory proteins include p25 (see
CC W50250), p22 (see W50241), Met-p22 (see W50247), p18 (see W50235),
CC Met-p18 (see W50237) and Met-p18-Het (see W50238). These proteins
CC can be produced by recombinant methods using claimed expression
CC vectors and host cells, and can be provided exogenously to target
CC cells for use in inhibiting HBV replication. Alternatively, a
CC nucleic acid construct that directs overexpression of an inhibitory
CC protein in target cells is used for the gene therapy of HBV
CC infection.
90 Sequence 183 AA;

FT misc_difference 178
 FT /note- "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT misc_difference 172..183
 FT /note- "amino acid residues 172-183, pref. 174-180,
 FT are deleted from the core protein in mutant
 FT polypeptides of the invention"
 PN W09700698-A1.
 PD 09-JAN-1997.
 PF 20-JUN-1996: U10602.
 PR 20-JUN-1995: US-017814.
 PA (GENO.) GEN. HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 97-087176/08.
 DR N-PSDB: T49594.
 PT New method for inhibiting the replication of hepadnaviruses -
 PT comprises introducing a mutant polypeptide with a mutated core
 PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
 PT B
 PS Claim 5: Page 55-56: 83pp: English.
 CC Alteration of the C-terminus of the core protein (W09044) of
 CC hepatitis B virus (HBV) or other hepadnavirus creates a mutant
 CC polypeptide capable of reducing replication of the wild-type virus
 CC by a dominant negative mechanism. The inhibitory effect is
 CC species-specific and is achieved by deletion of a few C-terminal
 CC amino acids from the core protein, and/or by joining the core
 CC protein to a hepadnavirus surface protein (see also W09045),
 CC creating a core-surface fusion. Such mutant polypeptides (see
 CC also W09046-50) can be used to treat hepadnavirus, e.g. HBV,
 CC infection.
 SQ Sequence 183 AA:

Query Match 45.4%; Score 1353; DB 20; Length 183;
 Best Local Similarity 98.9%; Pred. No. 1.04e-99;
 Matches 181; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 mdidpykefatvellsfspdffsvrvdlldtaaalrydalespncshphltalrga1 60
 |||
 Qy 1 MDIDPYKEFATVELLSFSPDFFSVRVLDLDTASALYREALSPENCSPHHTLRQA1L 60
 |||
 Db 61 cswgelmrlatwgvnledpsardlvsvyvtnglkrqlwfhiscrltfgrevieylv 120
 |||
 Qy 61 CWSGELMLATWGVNLEDPSARDLVSVYVTNMGKRLQFLWFIHISCTFGREVIEYLV 120
 |||
 Db 121 sfgywltppayrpnapilslpeltvrrgrsprrrrrpsrrrrsgare 180
 |||
 Qy 121 SFGWLTTPPAYRPNAPILSLPELTVVRGRSPRRRTPSPRRRSQSRRRSQSRE 180
 |||
 Db 181 sqc 183
 |||
 Qy 181 SQC 183

RESULT 9
 ID R05635 standard; protein; 183 AA.
 AC R05635;
 DT 30-OCT-1990 (first entry)
 DE Hepatitis B antigen.
 KM Hepatitis B virus; vaccine; HBV; ds.
 OS Synthetic.
 PN EP-374869-A.
 PD 27-JUN-1990.
 PF 1-JAN-1989: 123526.
 PR 22-DEC-1978: GB-049907.
 PR 27-DEC-1978: GB-050039.
 PR 01-NOV-1979: GB-037910.
 PA (BIOJ) Biogen Inc.
 PI Murray K, Schaller HE;
 DR WPI: 90-195067/26.
 DR N-PSDB: 004799
 PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
 PT used in detection of infection and in vaccine prodn.

PS Disclosure; 4pp: English.
 CC Polypeptide fragments encoded by the DNA sequence are antigenic for
 CC HBV and may be used as a vaccine or in detection. Peptides may be
 CC cultured in a suitable bacterial host such as E.coli.
 CC Fragments of the sequence are also claimed as being antigenically
 CC useful.
 SQ Sequence 183 AA:

Query Match 45.3%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 2.24e-99;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdidpykefatvellsfspdffsvrvdlldtaaalrydalespncshphltalrga1 60
 |||
 Qy 1 MDIDPYKEFATVELLSFSPDFFSVRVLDLDTASALYREALSPENCSPHHTLRQA1L 60
 |||
 Db 61 cswgelmrlatwgvnledpsardlvsvyvtnglkrqlwfhiscrltfgrevieylv 120
 |||
 Qy 61 CWSGELMLATWGVNLEDPSARDLVSVYVTNMGKRLQFLWFIHISCTFGREVIEYLV 120
 |||
 Db 121 sfgywltppayrpnapilslpeltvrrgrsprrrrrpsrrrrsgare 180
 |||
 Qy 121 SFGWLTTPPAYRPNAPILSLPELTVVRGRSPRRRTPSPRRRSQSRRRSQSRE 180
 |||
 Db 181 sqc 183
 |||
 Qy 181 SQC 183

RESULT 10
 ID P00041 standard; protein; 183 AA.
 AC P00041;
 DT 14-OCT-1992 (first entry)
 DE Sequence of core antigen.
 KM Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
 OS Hepatitis B virus.
 PN EP-13828-A.
 PD 06-AUG-1980.
 PF 21-DEC-1978: 303017.
 PR 22-DEC-1978: GB-049907.
 PR 27-DEC-1978: GB-050039.
 PR 01-NOV-1979: GB-037910.
 PA (BIOJ) BIOGEN NV.
 PI Murray K, Schaller HE;
 DR WPI: 80-57268C/33.
 DR N-PSDB: N00003.
 PT Recombinant DNA coding for polypeptide - have specificity of
 PT hepatitis B viral antigens in detection or antibody stimulation
 PS Example: Figs 3-4; 43pp: English.
 CC Human serum from a single HbsAg positive, HBeAg positive donor
 CC (serotype adym) was used to prep. a DNA-contg. pellet which was
 CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
 CC The labelled DNA was then extracted with phenol from the resulting
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
 CC It was then cloned in plasmid pBR322 which was used to transform E.
 CC coli. Micro-organisms prep. by the processes are deposited at the
 CC HBV-Kpn I dg: Tetr Amps HBV+.

SQ Sequence 183 AA:

Query Match 45.3%; Score 1349; DB 5; Length 183;
 Best Local Similarity 96.7%; Pred. No. 2.24e-99;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdidpykefatvellsfspdffsvrvdlldtaaalrydalespncshphltalrga1 60
 |||
 Qy 1 MDIDPYKEFATVELLSFSPDFFSVRVLDLDTASALYREALSPENCSPHHTLRQA1L 60
 |||
 Db 61 cswgelmrlatwgvnledpsardlvsvyvtnglkrqlwfhiscrltfgrevieylv 120
 |||
 Qy 61 CWSGELMLATWGVNLEDPSARDLVSVYVTNMGKRLQFLWFIHISCTFGREVIEYLV 120
 |||
 Db 121 sfgywltppayrpnapilslpeltvrrgrsprrrrrpsrrrrsgare 180
 |||

QY 121 SEGVWIRTPAPRPAPNAPILSTLPTETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
 Db 181 sgc 183
 111
 QY 181 SOC 183

RESULT 11
 ID P00004 standard; Protein; 184 AA.
 AC P00004;
 DT 14-OCT-1992 (first entry)
 DE Sequence of core antigen
 KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
 OS Hepatitis B virus.
 PN EP--13828-A.
 PD 06-AUG-1980.
 PF 21-DEC-1979: 303017.
 PR 22-DEC-1978: GB-049907.
 PR 27-DEC-1978: GB-050039.
 PR 01-NOV-1979: GB-037910.
 PA (BIOJ) BIOGEN NV.
 PI Murray K, Schaller HE;
 DR WPI: 80-57268C/33.
 N-PSDB: N00002.
 PT Recombinant DNA coding for polypeptide - have specificity of
 PT hepatitis B viral antigens in detection or antibody stimulation
 PS Claim 13: Page 40: 43pp; English.
 CC Human serum from a single HBSAg positive, HBeAg positive donor
 CC (serotype adym) was used to prep. a DNA-contg. pellet which was
 CC labeled with 3H or 32P as described by P. M. Kaplan et al (1973).
 CC The labeled DNA was then extracted with phenol from the resulting
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
 CC It was then cloned in plasmid pBR322 which was used to transform E.
 CC coli. Micro-organisms prepd. by the processes are deposited at the
 CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg;
 CC HBV-Kpn I dc; Tetr Amps HBV+.
 SQ Sequence 184 AA;

Query Match 44.9%; Score 1338; DB 5; Length 184;
 Best Local Similarity 96.2%; Pred. No. 1.85e-98;
 Matches 177; Conservative 5; Mismatches 1; Indels 1; Gaps 1;
 Db 1 mdidpykefgatvellsifpsdffpsvcrdlldtaaalyydaesphcsphtalrgal 60
 QY 1 MDIDPYKEFGATVELLSIFPSDFPSV-RDLDTASALYREALSPHCSPHTALRQAI 59
 Db 61 lcgwdimlatwgtntledpasrdlvsvyvnthvglkfgrllwfhlscltfgretvleyl 120
 QY 60 LCGWELMTLATWGVNLDPASRDLVSVYNTVMGLKFRQLMFHISCLTFGRETVLEYL 119
 Db 121 vsfgvwlrtppayrpnapilslptetvrrrgsrprrrpsprrrrsqsprrrrsqr 180
 QY 120 VSFGVWIRTPAPRPAPNAPILSTLPTETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 179
 Db 181 esgc 184
 1111
 QY 180 ESOC 183

RESULT 12
 ID R40806 standard; Protein; 196 AA.
 AC R40806;
 DT 16-FEB-1994 (first entry)
 DE Hepatitis B core / PV-1 fusion.
 KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
 OS Synthetic.
 FH key
 FT region
 FT protein
 FT J05192170-A.
 Location/Qualifiers
 7..13
 /label= PV-1
 14..196
 /label= HBC

PD 03-AUG-1993.
 PF 24-SEP-1991: 243800.
 PR 24-SEP-1991: JP-243800.
 PA (NISM) NISSHIN OIL MILLS LTD.
 DR WPI: 93-277479/35.
 DR N-PSDB: 047736.
 PT Recombinant plasmid for high immunogenicity virus - contains
 PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
 PT virus and exotic genes
 PS Disclosure; Fig 7; 12pp; Japanese.
 CC The hepatitis B core gene is recombinant with PV-1 DNA
 CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
 CC a vaccine. NB: Sequence is difficult to read in the original
 CC specification.
 SQ Sequence 196 AA;

Query Match 44.9%; Score 1336; DB 8; Length 196;
 Best Local Similarity 95.1%; Pred. No. 2.72e-98;
 Matches 175; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 Db 13 andidpykefgasvellsifpsdffpsirldldtasalyydaesphcsphtalrgal 72
 QY 213 SMDIDPYKEFGATVELLSIFPSDFPSVRLDTASALYREALSPHCSPHTALRQAI 272
 Db 73 lcgwdimlatwgtntledpasrdlvsvyvnthvglkfgrllwfhlscltfgretvleyl 132
 QY 273 LCGWELMTLATWGVNLDPASRDLVSVYNTVMGLKFRQLMFHISCLTFGRETVLEYL 132
 Db 133 vsfgvwlrtppayrpnapilslptetvrrrgsrprrrpsprrrrsqsprrrrsqr 192
 QY 333 VSFGVWIRTPAPRPAPNAPILSTLPTETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 392
 Db 193 esgc 196
 1111
 QY 393 ESOC 396

RESULT 13
 ID R40808 standard; Protein; 208 AA.
 AC R40808;
 DT 16-FEB-1994 (first entry)
 DE Hepatitis B core / PV-1 / IL-1 Fusion.
 KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
 OS Synthetic.
 FH key
 FT region
 FT region
 FT region
 FT protein
 FT J05192170-A.
 PN 03-AUG-1993.
 PD 24-SEP-1991: 243800.
 PR 24-SEP-1991: JP-243800.
 PA (NISM) NISSHIN OIL MILLS LTD.
 DR WPI: 93-277479/35.
 DR N-PSDB: 047738.
 PT Recombinant plasmid for high immunogenicity virus - contains
 PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
 PT virus and exotic genes
 PS Disclosure; Fig 9; 12pp; Japanese.
 CC The hepatitis B core gene is recombinant with PV-1 DNA
 CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
 CC a vaccine. NB: Sequence is difficult to read in the original
 CC specification.
 SQ Sequence 208 AA;

Query Match 44.9%; Score 1337; DB 8; Length 208;
 Best Local Similarity 94.6%; Pred. No. 2.24e-98;
 Matches 175; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 Db 24 igmdidpykefgasvellsifpsdffpsirldldtasalyydaesphcsphtalrga 83

```

QY 212 LMSDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCHHTALRQA 271
DB 84 Cwgelmlatwgsnledpasrelvsvyvmngklrkqlwfhiscitfgretvley 143
QY 272 ILWGELMTLATWGVNLEDPASRDVSVYVNTNMGKLFROLWFHISCLTFRGRTVLEY 331
DB 144 lvsfgwlrtpayrppnapilslipetvvrirgsprrrrsqsprrrrsqs 203
QY 332 LVSGWIRTPPAYRPPNAPILSLIPETTVVRRGRSPRRRTSPRRRSQSPRRRSQS 391
DB 204 resgc 208
QY 392 RESQC 396

```

```

RESULT 14
ID P40311 standard: Protein: 183 AA.
AC P40311;
DT 18-AUG-1992 (first entry)
DE Hepatitis virus core antigen.
KM HBcAg; vaccine; diagnosis; HBV infection.
OS Hepatitis b virus.
PN J59074985-A.
PD 27-APR-1984.
PF 19-OCT-1982; 183432.
PR 19-OCT-1982; JP-183432.
PA (TAKE ) TAKEDA CHEMICAL IND KK.
DR WPI: 84-143231/23.
PT DNA used in prevention of infections by hepatitis virus B -
PT comprises structural gene of hepatitis virus adr B surface antigen
PT coding gene and at least 1 virus core antigen structural gene.
PS Disclosure: Fig 3; 13p; Japanese.
CC The sequence is that of hepatitis virus core antigen (HBcAg). It
CC can be used as a vaccine for the prevention of infections by
CC hepatitis B virus (HBV) and also in the diagnosis of early stages
CC of HBV infection. See also P40310.
SQ Sequence 183 AA;

```

```

Query Match 44.88; Score 1334; DB 4; Length 183;
Best Local Similarity 95.68; Pred. No. 3.99e-98;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdldpykefgasvellsfjpsdfpsirdlldtasalyrealsepchphthalrqa1 60
QY 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCHHTALRQA1L 60
DB 61 Cwgelmlatwgsnledpasrelvsvyvmngklrkqlwfhiscitfgretvley1v 120
QY 61 Cwgelmlatwgsnledpasrelvsvyvmngklrkqlwfhiscitfgretvley1v 120
DB 121 sfvgwlrtpayrppnapilslipetvvrirgsprrrrsqsprrrrsqsre 180
QY 121 SFGWIRTPPAYRPPNAPILSLIPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSRE 180
DB 181 sgc 183
QY 181 SOC 183

```

```

RESULT 15
ID R40805 standard: Protein: 183 AA.
AC R40805;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core protein.
KM Hepatitis B: core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Hepatitis B virus.
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISW ) NISSHIN OIL MILLS LTD.
DR WPI: 93-277479/35.

```

```

DR N-PSDB; Q47735
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure: Fig 6; 12p; Japanese.
CC The hepatitis B core gene is recombined with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.
SQ Sequence 183 AA;

```

```

Query Match 44.88; Score 1334; DB 8; Length 183;
Best Local Similarity 95.68; Pred. No. 3.99e-98;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdldpykefgasvellsfjpsdfpsirdlldtasalyrealsepchphthalrqa1 60
QY 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCHHTALRQA1L 60
DB 61 Cwgelmlatwgsnledpasrelvsvyvmngklrkqlwfhiscitfgretvley1v 120
QY 61 Cwgelmlatwgsnledpasrelvsvyvmngklrkqlwfhiscitfgretvley1v 120
DB 121 sfvgwlrtpayrppnapilslipetvvrirgsprrrrsqsprrrrsqsre 180
QY 121 SFGWIRTPPAYRPPNAPILSLIPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSRE 180
DB 181 sgc 183
QY 181 SOC 183

```

```

Search completed: Thu Dec 16 12:47:00 1999
Job time : 202 secs.

```

 NW50E161
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_p protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Dec 16 12:47:19 1999; Maspar time 18.46 Seconds
 Tabular output not generated. 859,536 Million cell updates/sec

Title: >US38183-1-38183
 Description: (1-396) from us38183-1-38183.pep
 Perfect Score: 2977
 Sequence: 1 MDIDPYKEGATVELLSFLP.....RRRRSQSPRRRSQSRSSOC 396

Scoring table: PAM 150
 Gap 11

Searched: 122810 segs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: p160
 1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 46.943; Variance 114.459; scale 0.410

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	46.2	212	1	NKVLAH e antigen precursor / 1.92e-206	
2	1374	46.2	212	2	S53211 e antigen precursor / 2.85e-206	
3	1374	46.2	212	2	S53204 e antigen precursor / 2.85e-206	
4	1373	46.1	183	2	S53207 core antigen - hepati / 4.24e-206	
5	1373	46.1	212	2	S20750 e antigen precursor / 4.24e-206	
6	1370	46.0	183	2	S53270 core antigen - hepati / 1.39e-205	
7	1370	46.0	212	2	S53216 e antigen precursor / 1.39e-205	
8	1370	46.0	212	2	S53272 e antigen precursor / 1.39e-205	
9	1369	46.0	212	2	S53163 e antigen precursor / 2.07e-205	
10	1369	46.0	212	2	S20746 e antigen precursor / 2.07e-205	
11	1369	46.0	212	2	S53225 e antigen precursor / 2.07e-205	
12	1368	46.0	212	2	S53198 e antigen precursor / 3.08e-205	
13	1366	45.9	212	2	S53223 e antigen precursor / 6.81e-205	
14	1366	45.9	212	2	S53274 e antigen precursor / 6.81e-205	
15	1364	45.8	183	2	S53169 core antigen - hepati / 1.51e-204	
16	1363	45.8	183	2	S53247 core antigen - hepati / 2.24e-204	
17	1363	45.8	212	2	S53202 e antigen precursor / 2.24e-204	
18	1363	45.8	212	2	S53159 e antigen precursor / 2.24e-204	
19	1363	45.8	212	2	S53200 e antigen precursor / 2.24e-204	
20	1362	45.8	212	2	S53251 e antigen precursor / 3.33e-204	
21	1360	45.7	212	2	S53204 e antigen precursor / 3.33e-204	
22	1360	45.7	212	2	S53227 e antigen precursor / 7.37e-204	
23	1358	45.6	211	1	NKVLAI e antigen precursor / 1.63e-203	

24	1359	45.6	212	2	S53253 e antigen precursor / 1.10e-203
25	1357	45.6	212	2	S53281 e antigen precursor / 2.42e-203
26	1355	45.5	212	1	NKVLBH e antigen precursor / 5.35e-203
27	1355	45.5	212	2	S53242 e antigen precursor / 5.35e-203
28	1351	45.4	183	2	S53129 core antigen - hepati / 2.62e-202
29	1351	45.4	212	2	S53240 e antigen precursor / 2.62e-202
30	1349	45.3	183	1	NKVLAI core antigen - hepati / 5.78e-202
31	1349	45.3	183	2	S53181 core antigen - hepati / 5.78e-202
32	1348	45.3	183	2	S53232 core antigen - hepati / 8.60e-202
33	1348	45.3	183	2	S53260 core antigen - hepati / 8.60e-202
34	1345	45.2	183	2	S53152 core antigen - hepati / 2.83e-201
35	1343	45.1	183	2	S53214 core antigen - hepati / 6.25e-201
36	1342	45.1	183	2	S53140 core antigen - hepati / 9.28e-201
37	1344	45.1	212	2	S53229 e antigen precursor / 4.20e-201
38	1341	45.0	183	2	S53267 core antigen - hepati / 1.38e-200
39	1341	45.0	183	2	S53189 core antigen - hepati / 1.38e-200
40	1339	45.0	183	2	S53137 core antigen - hepati / 3.05e-200
41	1341	45.0	212	2	S53238 e antigen precursor / 1.38e-200
42	1340	45.0	212	2	S53257 e antigen precursor / 2.05e-200
43	1339	45.0	212	2	S23651 e antigen precursor / 3.05e-200
44	1338	44.9	212	2	S53236 e antigen precursor / 4.54e-200
45	1337	44.9	212	1	NKVLJ2 e antigen precursor / 6.75e-200

ALIGNMENTS

RESULT 1
 ENTRY NKVLAH #type complete
 TITLE e antigen precursor / core antigen - hepatitis B virus
 (subtype ayw4, isolate hb321 and others)
 ALTERNATE_NAMES HBe antigen precursor / HBe antigen; pre-C/C antigen
 CONTRAINS core antigen; e antigen
 ORGANISM #formal_name hepatitis B virus; HBV
 #variety subtype ayw4, isolate hb321; isolate patient Ferracuti/83;
 isolate patient Castag/83; isolate patient Sanna/84;
 isolate patient Licheri-1/85; isolate patient Flore-1/86;
 isolate patient Licheri/83
 18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change
 10-Oct-1997

DATE S47405; S53191; S53209; S53234; S53264; S53249; S53262;
 S53277; A03711

ACCESSIONS S47404
 REFERENCE Plucienniczak, A.
 #authors Submitted to the EMBL Data Library, August 1994
 #submission Molecular cloning and sequencing of two complete genomes of
 #description Polish isolates of human hepatitis B virus.
 #accession S47405

#molecule_type DNA
 #residues 1-212 #label PUJ
 #cross-references EMBL:235716; NID:g527435; PID:g527437
 #experimental_source subtype ayw4, isolate hb321
 S53112
 REFERENCE Lai, M.E.; Mazzeo, A.P.; Porru, A.; Balestrieri, A.
 #authors Submitted to the EMBL Data Library, March 1995
 #submission
 #description
 #accession S53191

#molecule_type DNA
 #residues 1-212 #label LAI
 #cross-references EMBL:X85283; NID:g736088; PID:g736090
 #experimental_source isolate patient Ferracuti/83
 #accession S53209

#molecule_type DNA
 #residues 1-212 #label LA2
 #cross-references EMBL:X85290; NID:g736114; PID:g736116
 #experimental_source isolate patient Castag/83
 #accession S53234

#molecule_type DNA
 #residues 1-212 #label LA3
 #cross-references EMBL:X85300; NID:g736150; PID:g736152
 #experimental_source isolate patient Sanna/84
 #accession S53264

#molecule_type DNA
 #residues 1-212 #label LA4
 #cross-references EMBL:X85313; NID:g736194; PID:g736196

##experimental_source isolate patient Licheri-1'85
#accession S53249
##molecule_type DNA
##residues 1-212 ##label LA5
##cross-references EMBL:X85306; NID:g736172; PID:g736174
##experimental_source isolate patient Flore-1'86
#accession S53262
##molecule_type DNA
##residues 1-212 ##label LA6
##cross-references EMBL:X85312; NID:g736191; PID:g736193
##experimental_source isolate patient Licheri'83
#accession S53277
##molecule_type DNA
##residues 30-212 ##label LA7
##cross-references EMBL:X85317; NID:g736211; PID:g736214
##experimental_source patient Gordo-2'86
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be processed

REFERENCE A93214
#authors Galibert, F.; Mandart, E.; Flitoussi, F.; Tiohais, P.;
#journal Nature (1979) 281:646-650
#title Nucleotide sequence of the hepatitis B virus genome (subtype
ayw) in F. coli.
#cross-references M01D:81012091
#accession A03711
##molecule_type DNA
##residues 1-212 ##label GAL
##cross-references GB:J02203; NID:g329640; PID:g329642
##experimental_source subtype ayw

GENETICS C
#gene #superfamily hepatitis B virus core antigen
#KEYWORDS alternative initiators; core protein
FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24350 #checksum 782

Query Match 46.2%; Score 1375; DB 1; Length 212;
Best Local Similarity 99.5%; Pred. No. 1,926-206;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFSDFFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 213 SMDIDPKKEGATVELLSFSDFFPSVRDLDTASALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNEDPASRDLYSVYNTNGKLFROLWFHISCLTFGRTVIEYL 148
QY 273 LCMGELMTLATWGVNEDPASRDLYSVYNTNGKLFROLWFHISCLTFGRTVIEYL 332

Db 149 VSEGVWIRTPAPYPPNAPILSTPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
QY 333 VSEGVWIRTPAPYPPNAPILSTPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY 393 ESOC 396

RESULT 2
ENTRY S53211 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Castag-1'85 and others)
ALTERNATE_NAMES HBC antigen; HBe antigen precursor / HBC antigen; pre-C/C
antigen
CONTAINS core antigen; e antigen
ORGANISM #formal name hepatitis B virus, HBV
#isolate patient Castag-1'85; isolate patient Ferracuti-1'89
#variety 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

08-Sep-1997
ACCESSIONS S53211; S53197
REFERENCE S53112
#authors Lai, M.E.; Mazzeoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53211
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85291; NID:g736117; PID:g736119
##experimental_source isolate patient Castag-1'85
#accession S53197
##molecule_type DNA
##residues 30-212 ##label LAN
##cross-references EMBL:X85284; NID:g736095; PID:g736098
##experimental_source isolate patient Ferracuti-1'89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS C
#gene #superfamily hepatitis B virus core antigen
#KEYWORDS alternative initiators; core protein
FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24336 #checksum 746

Query Match 46.2%; Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 2,856-206;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFSDFFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 213 SMDIDPKKEGATVELLSFSDFFPSVRDLDTASALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNEDPASRDLYSVYNTNGKLFROLWFHISCLTFGRTVIEYL 148
QY 273 LCMGELMTLATWGVNEDPASRDLYSVYNTNGKLFROLWFHISCLTFGRTVIEYL 332

Db 149 VSEGVWIRTPAPYPPNAPILSTPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
QY 333 VSEGVWIRTPAPYPPNAPILSTPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY 393 ESOC 396

RESULT 3
ENTRY S32204 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1000)
ALTERNATE_NAMES HBe antigen precursor / HBe antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal name hepatitis B virus, HBV
#variety subtype ayw, patient C1000
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997

ACCESSIONS S32204
REFERENCE S32202
#authors Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.;
Gerok, W.; Rasenack, J.
#submission submitted to the EMBL Data Library, March 1993
#description Identification and sequence analysis of hepatitis B virus DNA
in immunological negative infection.
#accession S32204
##molecule_type DNA
##residues 1-212 ##label PRE
##cross-references EMBL:X72702; NID:g288927; PID:g288930
##experimental_source subtype ayw, patient C1000
#note due to a stop codon between the alternative initiators

the e antigen precursor cannot be produced

GENETICS

CLASSIFICATION C #superfamily hepatitis B virus core antigen

KEYWORDS core protein

FEATURE

1-29 #domain signal sequence #status predicted #label SIG
30-212 #product core antigen #status predicted #label CAG
30-178 #product e antigen #status predicted #label EAG
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24336 #checksum 770

Query Match 46.2%; Score 1374; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 2,856-206;

Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88

QY 213 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 148

QY 273 LCMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 332

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208

QY 333 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212

QY 393 ESOC 396

RESULT

ENTRY 4 S53207 #type complete

TITLE core antigen - hepatitis B virus (isolate patient

ALTERNATE_NAMES HBC antigen

CONTAINS core antigen

ORGANISM #formal_name hepatitis B virus, HBV

DATE #isolate_patient Castaa-2'87

08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

ACCESSIONS

REFERENCE S53207

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53207

##molecule_type DNA

##residues 1-183 ##label LAI

##cross-references EMBL:X85289; NID:g736110; PID:g736113

##experimental_source isolate patient Castaa-2'87

##note due to a stop codon between the alternative initiators

the e antigen precursor cannot be produced

GENETICS

CLASSIFICATION C #superfamily hepatitis B virus core antigen

KEYWORDS core protein

SUMMARY #length 183 #molecular-weight 21102 #checksum 2165

Query Match 46.1%; Score 1373; DB 2; Length 183;

Best Local Similarity 99.5%; Pred. No. 4,246-206;

Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 60

QY 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 60

Db 61 CMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 120

QY 61 CMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 120

Db 121 SFGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180

QY 121 SFGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180

Db 181 SOC 183

QY 181 SOC 183

RESULT

ENTRY 5 S20750 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES (subtype ayw, patient CI)

CONTAINS HBE antigen precursor / HBC antigen; pre-C/C antigen

ORGANISM core antigen; e antigen

DATE #variety #formal_name hepatitis B virus, HBV

08-Sep-1997 subtype ayw, patient CI

20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change

ACCESSIONS

REFERENCE S20750

#authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.;

#submission submitted to the EMBL Data Library, March 1992

#description Sequence analysis of HBV genomes isolated from patients with

#accession S20750

##molecule_type DNA

##residues 1-212 ##label LAI

##cross-references EMBL:X65258; NID:g59434; PID:g59436

##experimental_source subtype ayw, patient CI

GENETICS

CLASSIFICATION C #superfamily hepatitis B virus core antigen

KEYWORDS alternative initiators; core protein

FEATURE

1-29 #domain signal sequence #status predicted #label SIG\

30-212 #product core antigen #status predicted #label CAG\

30-178 #product e antigen #status predicted #label EAG\

179-212 #domain carboxyl-terminal propeptide #link EAG #status

predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 147

Query Match 46.1%; Score 1373; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 4,246-206;

Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88

QY 213 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 148

QY 273 LCMGELMTLATWGVNLEDPASRDVLVSYVTNMGKFRQLMFIHISCLTREGRETVIEL 332

Db 149 VAFGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208

QY 333 VAFGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212

QY 393 ESOC 396

RESULT

ENTRY 6 S53207 #type complete

TITLE core antigen - hepatitis B virus (isolate patient

ALTERNATE_NAMES Licheri-2'87)

CONTAINS HBC antigen

ORGANISM #formal_name hepatitis B virus, HBV

DATE #variety #isolate_patient Licheri-2'87

08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

```

08-Sep-1997
ACCESSIONS      S53370
REFERENCE        S53112
#authors        Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submissions    submitted to the EMBL Data Library, March 1995
#accession      S53370
##molecule_type DNA
##residues      1-183 ##label LAI
##cross-references EMBL:X85314; NID:g736201; PID:g736204
##experimental_source isolate patient Licheri-2/87
#note           due to a stop codon between the alternative initiators
                the e antigen precursor cannot be produced

GENETICS
#gene           C
CLASSIFICATION  #superfamily 'hepatitis B virus core antigen
KEYWORDS        core protein
SUMMARY         #length 183 #molecular-weight 21102 #checksum 2199

Query Match      46.0%; Score 1370; DB 2; Length 183;
Best Local Similarity 98.4%; Pred. No. 1.39e-205;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDYKEFGATVELLSLPDSDFPVSVDLDDTASALYRDALSPHCSPHHTALRQAIL 60
      |||||
Oy 1 MDIDYKEFGATVELLSLPDSDFPVSVDLDDTASALYREALSPHCSPHHTALRQAIL 60
      |||||

Db 61 CWDGLMTLATWGVNLDPASRDLYSVYNTNMGKFRQLMFIHISCLTFRGFEVIEYLV 120
      |||||
Oy 61 CWDGLMTLATWGVNLDPASRDLYSVYNTNMGKFRQLMFIHISCLTFRGFEVIEYLV 120
      |||||

Db 121 SFGVWIRTPPAYRPNNAILSTLPETTVVRRGRTPRRRTPSPRRRSQSPRRRSQSPRE 180
      |||||
Oy 121 SFGVWIRTPPAYRPNNAILSTLPETTVVRRGRTPRRRTPSPRRRSQSPRRRSQSPRE 180
      |||||

Db 181 SOC 183
      |||
Oy 181 SOC 183

RESULT 7
ENTRY      S53216 #type complete
TITLE      e antigen precursor / core antigen - hepatitis B virus
            (isolate patient Castag'3)
ALTERNATE_NAMES
CONTAINS   HBe antigen precursor / Hbc antigen; pre-C/C antigen
ORGANISM   core antigen; e antigen
            #formal_name hepatitis B virus; HBV
DATE       #variety isolate patient Castag'3
            #08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
            08-Sep-1997
ACCESSIONS S53216
REFERENCE   S53112
#authors   Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submissions submitted to the EMBL Data Library, March 1995
#accession S53216
##molecule_type DNA
##residues 1-212 ##label LAI
##cross-references EMBL:X85293; NID:g736124; PID:g736126
##experimental_source isolate patient Castag'3

GENETICS
#gene           C
CLASSIFICATION  #superfamily 'hepatitis B virus core antigen
KEYWORDS        alternative initiators; core protein
FEATURE         #domain signal sequence #status predicted #label SVE\
1-29            #product core antigen #status predicted #label CAG\
30-212          #product e antigen #status predicted #label EAG\
30-118          #domain carboxyl-terminal propeptide #link EAG #status
179-212        predicted #label ECP
SUMMARY         #length 212 #molecular-weight 24563 #checksum 752

Query Match      46.0%; Score 1370; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.39e-205;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

D	b		29	GMDIDPYKEFATYELLSEFLPSDFPSVRLDLDASALYRALESPEHCSPHNTALQAI	88
Oy			213	SMDDIPYKEFATYELLSEFLPSDFPSVRLDLDASALYRALESPEHCSPHNTALQAI	272
D	b		89	LWGEMLMTLATWGVNEDPASRDLVSVYNTNGKLFROLLMFIICLTFGRETVEYL	148
Oy			273	LWGEMLMTLATWGVNEDPASRDLVSVYNTNGKLFROLLMFIICLTFGRETVEYL	332
D	b		149	VSFGWITTPAYRPNPAPILSTLPETTVARRRGRSRPRRRRSOSPPRRRSOSR	208
Oy			333	VSFGWITTPAYRPNPAPILSTLPETTVARRRGRSRPRRRRSOSPPRRRSOSR	392
D	b		209	ESOC 212	
Oy			393	ESOC 396	
R	E	S	8		
ENTRY				#type complete	
TITLE				e antigen precursor / core antigen - hepatitis B virus (isolate patient Licheri-3'90)	
ALTERNATE_NAMES				HBe antigen precursor / Hbc antigen; pre-C/C antigen	
CONTAINS				Core antigen; e antigen	
ORGANISM				#formal_name hepatitis B virus, HBV	
DATE				isolate patient Licheri-3'90	
				08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change	
				08-Sep-1997	
ACCESSIONS				S53372	
REFERENCE				S53312	
#authors				Lai, M.E.; Mazzeoli, A.P.; Porru, A.; Balestrieri, A.	
#submitted				submitted to the EMBL Data Library, March 1995	
#accession				S53372	
#molecule_type				DNA	
#residues				1-212 #label LAI	
#cross-references				EMBL:X83515, NID:g736205, PID:g736207	
#experimental_source				Isolate patient Licheri-3'90	
GENETICS					
#gene				C	
CLASSIFICATION				#superfamily hepatitis B virus core antigen	
KEYWORDS				alternative initiators; core protein	
FEATURE					
1-29				#domain signal sequence #status predicted #label SIG\	
30-212				#product core antigen #status predicted #label CAG\	
30-178				#product e antigen #status predicted #label EAG\	
179-212				#domain carboxyl-terminal propeptide #link EAG #status	
				predicted #label ECP	
SUMMARY				#length 212 #molecular_weight 24366 #checksum 446	
Query Match				46.0%; Score 1370; DB 2; Length 212;	
Best Local Similarity				98.9%; Pred. No. 1,39e-205;	
Matches 182; Conservative				1; Mismatches 1; Indels 0; Gaps 0	
D	b		29	GMDIDPYKEFATYELLSEFLPSDFPSVRLDLDASALYRALESPEHCSPHNTALQAI	88
Oy			213	SMDDIPYKEFATYELLSEFLPSDFPSVRLDLDASALYRALESPEHCSPHNTALQAI	272
D	b		89	LWGEMLMTLATWGVNEDPASRDLVSVYNTNGKLFROLLMFIICLTFGRETVEYL	148
Oy			273	LWGEMLMTLATWGVNEDPASRDLVSVYNTNGKLFROLLMFIICLTFGRETVEYL	332
D	b		149	VSFGWITTPAYRPNPAPILSTLPETTVARRRGRSRPRRRRSOSPPRRRSOSR	208
Oy			333	VSFGWITTPAYRPNPAPILSTLPETTVARRRGRSRPRRRRSOSPPRRRSOSR	392
D	b		209	ESOC 212	
Oy			393	ESOC 396	
R	E	S	9		
ENTRY				S53163	
TITLE				#type complete	
				e antigen precursor / core antigen - hepatitis B virus	

ALTERNATE_NAMES	(isolate patient Vittorina'92)
CONTRAINS	HBe antigen precursor / Hbc antigen; pre-C/C antigen
ORGANISM	core antigen; e antigen
variety	#formal_name hepatitis B virus, HBV
DATE	isolate patient Vittorina'92 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997
ACCESSIONS	S53163
REFERENCE	S53112
authors	Lai, M.-E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submission	submitted to the EMBL Data Library, March 1995
accession	S53163
##residues	##molecule_type DNA
##cross-references	1-212 ##label LAI
##experimental_source	##cross-references EMBL:X65556; NID:g736050; PID:g736052
	##experimental_source isolate patient Vittorina'92
GENETICS	
CLASSIFICATION	C
KEYWORDS	#superfamily hepatitis B virus core antigen
FEATURE	alternative initiators; core protein
1-29	
30-212	#domain signal sequence #status predicted #label SIG\
30-178	#product core antigen #status predicted #label CAG\
179-212	#product e antigen #status predicted #label EAG\ #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP
SUMMARY	#length 212 #molecular_weight 24322 #checksum 842
Query Match	46.0%; Score 1369; DB 2; Length 212;
Best Local Similarity	98.9%; Pred. No. 2,07e-205;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
Db 29 GMDIDPKYEFATVELSLFSPDFPSYRDLSTASLYRALESPPHCSPHTALROAI 88	
Oy 213 SMDIDPYKEFATVELSLFSPDFPSYRDLSTASLYRALESPPHCSPHTALROAI 272	
Db 89 LCMGELMTLATWGVNEDPASRDLYSVYNTNGALFRQLMFIHISCLTFEGRETVIEYL 148	
Oy 213 LCMGELMTLATWGVNEDPASRDLYSVYNTNGALFRQLMFIHISCLTFEGRETVIEYL 332	
Db 149 VSFQGWIRTPAYRPNPAPILSTPETTVARRGRSSPRRRRSSQSPRRRSOSQR 208	
Oy 333 VSFGWIRTPAYRPNPAPILSTPETTVARRGRSSPRRRRSSQSPRRRSOSQR 392	
Db 209 ESQC 212	
Oy 393 ESQC 396	
RESULT 10	
ENTRY	S20746 #type complete
TITLE	e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient C)
ALTERNATE_NAMES	HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS	Core antigen; e antigen
ORGANISM	#formal_name hepatitis B virus, HBV
variety	subtype ayw, patient C
DATE	20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997
ACCESSIONS	S20746
REFERENCE	S20745
authors	Lai, M.-E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
submission	submitted to the EMBL Data Library, March 1992
description	Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic liver disease.
accession	S20746
##molecule_type	DNA
##residues	1-212 ##label LAI
##cross-references	EMBL:X65557; NID:g59429; PID:g59431
##experimental_source	##experimental_source subtype ayw, patient C
GENETICS	
gene	C

CLASSIFICATION	#superfamily hepatitis B virus core antigen
KEYWORDS	alternative initiators; core protein
FEATURE	
1-29	#domain signal sequence #status predicted #label SIG\
30-212	#product core antigen #status predicted #label CAG\
30-118	#product e antigen #status predicted #label EAG\
179-212	#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP
SUMMARY	#length 212 #molecular-weight 24348 #checksum 506
Query Match	46.0%: Score 1369; DB 2; Length 212;
Best Local Similarity	98.4%: Pred. No. 2,07e-205;
Matches	181; Conservative 2; Mismatches 1; Indels 0; Gaps 0
Db	29 GMDIDPYKEFGATVELLSFLPSDFPSPYRDLDTASALYREALESPHCSPHHTALROAI 88
Qy	213 SMDIDPYKEFGATVELLSFLPSDFPSPYRDLDTASALYREALESPHCSPHHTALROAI 272
Db	89 LCMGELMTLAWGVNEDPASPRLVSVYVNTNGGLFROLMHFISCLTFGRRTVLEYL 148
Qy	273 LCMGELMTLAWGVNEDPASPRLVSVYVNTNGGLFROLMHFISCLTFGRRTVLEYL 332
Db	149 VSFGEWIRTPPYRPPNAPILSTLPETTVARRGRSPRRRTPSRRRRSQSR 208
Qy	333 VSFGEWIRTPPYRPPNAPILSTLPETTVARRGRSPRRRTPSRRRRSQSR 392
Db	209 ESQC 212
Qy	393 ESQC 396
RESULT	11
ENTRY	S53225 #type complete
TITLE	e antigen precursor / core antigen - hepatitis B virus
ALTERNATE_NAMES	(isolate patient Chigline-2/86)
CONTAINS	HBe antigen precursor / HBe antigen; pre-C/C antigen
ORGANISM	core antigen; e antigen
#variety	#formal.name hepatitis B virus, HBV
DATE	isolate patient Chigline-2/86
08-Jul-1995	#sequence_revision 03-Aug-1995 #text-change
08-Sep-1997	
ACCESSIONS	S53225
REFERENCE	S53112
#authors	Lai, M.E.; Marzotlen, A.P.; Portu, A.; Balesstrierl, A.
#submission	submitted to the EMBL Data Library, March 1995
#accession	S53225
##residues	1-212 #label LAI
##molecule_type	DNA
##cross-references	EMBL:X85396; NID:g736137; PID:g736139
##experimental_source	isolate patient Chigline-2/86
GENETICS	
#gene	C
CLASSIFICATION	#superfamily hepatitis B virus core antigen
KEYWORDS	alternative initiators; core protein
FEATURE	
1-29	#domain signal sequence #status predicted #label SIG\
30-212	#product core antigen #status predicted #label CAG\
30-118	#product e antigen #status predicted #label EAG\
179-212	#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP
SUMMARY	#length 212 #molecular-weight 24308 #checksum 655
Query Match	46.0%: Score 1369; DB 2; Length 212;
Best Local Similarity	98.4%: Pred. No. 2,07e-205;
Matches	181; Conservative 3; Mismatches 0; Indels 0; Gaps 0
Db	29 GMDIDPYKEFGATVELLSFLPSDFPSPYRDLDTASALYREALESPHCSPHHTALROAI 88
Qy	213 SMDIDPYKEFGATVELLSFLPSDFPSPYRDLDTASALYREALESPHCSPHHTALROAI 272
Db	89 LCMGELMTLAWGVNEDPASPRLVSVYVNTNGGLFROLMHFISCLTFGRRTVLEYL 148
Qy	273 LCMGELMTLAWGVNEDPASPRLVSVYVNTNGGLFROLMHFISCLTFGRRTVLEYL 332

Db	149	VSFGWITTPAYRPNAPILISTLPEITTVARRGRSPRRRTPPRRRRSSOPRRRRSQSR	208
OY	333	VSFGWITTPAYRPNAPILISTLPEITTVARRGRSPRRRTPPRRRRSSOPRRRRSQSR	392
Db	209	ESOC 212	
OY	393	ESOC 396	
RESULT	12		
ENTRY		S53198	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus (isolate patient Ferracuti-2/90)	
ALTERNATE_NAMES		HBe antigen precursor / HBe antigen; pre-C/C antigen	
CONTAINS		core antigen; e antigen	
ORGANISM		#formal_name hepatitis B virus; HBV	
DATE		isolate patient Ferracuti-2/90	
DATE		08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		S53198	
REFERENCE		S53198	
authors		Lai, M. E.; Mazzoleni, A. P.; Portu, A.; Balestrieri, A.	
submission		submitted to the EMBL Data Library, March 1995	
accession		S53198	
residues		#molecule_type DNA	
residues		1-212 #label IAI	
cross-references		EMBL:X85285; NID:q736099; PID:q736100	
experimental_source		isolate patient Ferracuti-2/90	
GENETICS			
gene		C	
CLASSIFICATION		#superfamily hepatitis B virus core antigen	
KEYWORDS		alternative initiators; core protein	
FEATURE			
1-29		#domain signal sequence #status predicted #label SIG\	
30-212		#product core antigen #status predicted #label CAG\	
30-178		#product e antigen #status predicted #label EAG\	
179-212		#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY		#length 212 #molecular_weight 24292 #checksum 593	
Query Match		46.0%; Score 1368; DB 2; Length 212;	
Best Local Similarity		98.4%; Pred. No. 3.08e-205;	
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;			
Db	29	GMDIDPYKEFGATVELSLFSPSAFFPSYRDLLDTPASALYREALSPHCSPHNTALRQAI	88
OY	213	SMDDIDPYKEFGATVELSLFSPDFPSYRDLLDTPASALYREALSPHCSPHNTALRQAI	272
Db	89	LCMGDLMLTAVGWGNEDPASRDLYSVYVNTNGKLFROLMLPHISCLTFGRITVEYL	148
OY	273	LCMGDLMLTAVGWGNEDPASRDLYSVYVNTNGKLFROLMLPHISCLTFGRITVEYL	332
Db	149	VSFGWITTPAYRPNAPILISTLPEITTVARRGRSPRRRTPPRRRRSSOPRRRRSQSR	208
OY	333	VSFGWITTPAYRPNAPILISTLPEITTVARRGRSPRRRTPPRRRRSSOPRRRRSQSR	392
Db	209	ESOC 212	
OY	393	ESOC 396	
RESULT	13		
ENTRY		S53223	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus (isolate patient Chigline-1/85)	
ALTERNATE_NAMES		HBe antigen precursor / HBe antigen; pre-C/C antigen	
CONTAINS		core antigen; e antigen	
ORGANISM		#formal_name hepatitis B virus; HBV	
DATE		isolate patient Chigline-1/85	
DATE		08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		S53223	

REFERENCE	ENTRY	TYPE	COMPLETE
S53112	La1, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.		
#authors	La1, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.		
#submitted	submitted to the EMBL Data Library, March 1995		
#accession	S53223		
##molecule_type	DNA		
##residues	1-212		
##cross-references	EMBL:X85295; NID:g736134; PID:g736136		
##experimental_source	isolate patient Chaigne-1-85		
GENETICS			
#gene	C		
CLASSIFICATION	#superfamily hepatitis B virus core antigen		
KEYWORDS	alternative initiators; core protein		
FEATURE			
1-29			
30-212	#domain signal sequence #status predicted #label SIG\		
30-178	#product core antigen #status predicted #label CAG\		
179-212	#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP		
SUMMARY	#length 212 #molecular-weight 24364 #checksum 1123		
Query Match	45.9%; Score 1366; DB 2; Length 212;		
Best Local Similarity	97.8%; Pident. No. 6, 816-205;		
Matches 180; Conservative	3; Mismatches 1; Indels 0; Gaps 0		
Db	29 GMDIDPKRGATVLLSPFSPDFPSVVDLDTASALYREALSPHCSPHHTALROAI 88		
Oy	213 SMDIDPKRGATVLLSPFSPDFPSVVDLDTASALYREALSPHCSPHHTALROAI 272		
Db	89 LCMGDLMSLATWGVNLDPIDPSIDLVSVYNTNMGKFRQLMFIHSCITFGRETIVELY 148		
Oy	273 LCMGELMTLATWGVNLDPASNDLVSVYNTNMGKFRQLMFIHSCITFGRETIVELY 332		
Db	149 VSEGVWIRTPAPRPPAPLSTLPETTYVRRKGSPPRRKTPSPRRRSQSPRRRSQSR 208		
Oy	333 VSEGVWIRTPAPRPPAPLSTLPETTYVRRKGSPPRRKTPSPRRRSQSPRRRSQSR 392		
Db	209 ESOC 212		
Oy	393 ESOC 396		
RESULT 14			
ENTRY	S53274	#type complete	
TITLE	e antigen precursor / core antigen - hepatitis B virus		
ALTERNATE_NAMES	(isolate patient Giordano/84)		
CONTAINS	Hbe antigen precursor / Hbc antigen; pre-C/C antigen		
ORGANISM	core antigen; e antigen		
ONCOTYPE	#formal_name hepatitis B virus; HBV		
##variety	isolate patient Giordano/84		
DATE	08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997		
ACCESSIONS	S53274		
REFERENCE	S53112		
#authors	La1, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.		
#submission	submitted to the EMBL Data Library, March 1995		
#accession	S53274		
##molecule_type	DNA		
##residues	1-212		
##cross-references	EMBL:X85316; NID:g736208; PID:g736210		
##experimental_source	isolate patient Giordano/84		
GENETICS			
#gene	C		
CLASSIFICATION	#superfamily hepatitis B virus core antigen		
KEYWORDS	alternative initiators; core protein		
FEATURE			
1-29			
30-212	#domain signal sequence #status predicted #label SIG\		
30-178	#product core antigen #status predicted #label CAG\		
179-212	#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP		
SUMMARY	#length 212 #molecular-weight 24376 #checksum 390		
Query Match	45.9%; Score 1366; DB 2; Length 212;		

Best Local Similarity 98.9%: Pred. No. 6.81e-205;
Matches 182: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
OY 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
Db 89 LCWGEMLMTATWGVNLEDPSARDLVSYVNTNMGKFRQLMFHISCLTFGRETIVIEYL 148
OY 273 LCWGEMLMTATWGVNLEDPSARDLVSYVNTNMGKFRQLMFHISCLTFGRETIVIEYL 332
Db 149 VSFGWIRTPPAYRPPNAPILTLPTETVVRRRGRSPRRRTSPRRRSQSPRRRSQSR 208
OY 333 VSFGWIRTPPAYRPPNAPILTLPTETVVRRRGRSPRRRTSPRRRSQSPRRRSQSR 392
Db 209 ESOC 212
OY 393 ESOC 396

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu '89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#varietal_name isolate patient Muresu '89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53169
REFERENCE S53112
#authors Lai, M.E.; Mazzeoni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
#molecule_type DNA

##residues 1-183 ##label LAI
##cross-references EMBL:X85275; NID:g736057; PID:g736060
##experimental_source isolate patient Muresu '89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular-weight 21088 #checksum 1849

Query Match 45.8%: Score 1364; DB 2; Length 183;
Best Local Similarity 98.4%: Pred. No. 1.51e-204;
Matches 180: Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
OY 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
Db 61 CWGEMLMTATWGVNLEDPSARDLVSYVNTNMGKFRQLMFHISCLTFGRETIVIEYL 120
OY 61 CWGEMLMTATWGVNLEDPSARDLVSYVNTNMGKFRQLMFHISCLTFGRETIVIEYL 120
Db 121 SFGWIRTPPAYRPPNAPILTLPTETVVRRRGRSPRRRTSPRRRSQSPRRRSQSR 180
OY 121 SFGWIRTPPAYRPPNAPILTLPTETVVRRRGRSPRRRTSPRRRSQSPRRRSQSR 180
Db 181 SOC 183
OY 181 SOC 183

Search completed: Thu Dec 16 12:48:36 1999
Job time : 77 secs.

(TM)

Distribution rights by Oxford Molecular Ltd

t generated.

MDIDE

Зап 11

Listing first 45

1:swissprot

mean 48.173; variance 102.286; scale 0.471

ved by analysis of the total score distribution.

5.745 1

Best Local

Best Local

Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 60
QY 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 60
Db 61 CMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 120
QY 61 CMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 120
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRSOSRE 180
QY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRSOSRE 180
Db 181 SOC 183
QY 181 SOC 183
QY 181 SOC 183

RESULT 2
ID CORA_HPBVA STANDARD; PRT; 211 AA.
AC P24023;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN ALPHA).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 90266476.
RA TONG S., LI J., VIRIVINSKI L., TREPO C.;
RT "Active hepatitis B virus replication in the presence of anti-HBe is
RT associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC EMBL: M32138; G495034; ALT_SEQ.
DR PIR: A34773; NKVLAI.
DR PFAM: PF00906; Hepatitis_core; 1.
DR KW CORE PROTEIN; Repeat
FT REPEAT 183 190
FT REPEAT 198 206
SQ SEQUENCE 211 AA; 24208 MW; 0924CE5E CRC32;
```

Query Match 45.6%; Score 1358; DB 1; Length 211;
Best Local Similarity 95.2%; Pred. No. 2.35e-233;
Matches 179; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Db 24 LGWLDMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 83
QY 209 LGWLDMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 268
Db 84 ROAILCWGDLITLSTWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 143
QY 269 ROAILCWGDLITLSTWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 328
Db 144 IEVLVSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRR 203
QY 329 IEVLVSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRR 388
Db 204 TOSRESOC 211
QY 389 SOSRESOC 396
```

RESULT 3
ID CORA_HPBVA STANDARD; PRT; 183 AA.

```
AC P03147;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADYN).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 81012115.
RA PASEK M., GOTO T., GILBERT W., ZINK B., SCHALLER H., MCKAY P.,
RA LEADBETTER G., MURRAY K.;
RT "Hepatitis B virus genes and their expression in E. coli.";
RL NATURE 282:575-579(1979).
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC EMBL: J02202; G329638; -.
DR EMBL: A08967; G411874; -.
DR PIR: B93217; NKVLAI2
DR PFAM: PF00906; Hepatitis_core; 1.
DR KW CORE PROTEIN; Repeat
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;
```

Query Match 45.3%; Score 1349; DB 1; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.41e-231;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 60
QY 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTAL 60
Db 61 CMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 120
QY 61 CMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEIYLV 120
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRSOSRE 180
QY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRSOSRE 180
Db 181 SOC 183
QY 181 SOC 183
QY 181 SOC 183

RESULT 4
ID CORA_HPBVA STANDARD; PRT; 183 AA.
AC P17392;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN OKINAWA/P0DW282).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
RA MIYAKAWA Y., NAYUMI M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.";
```


RL J. GEN. VIROL. 69:2575-2583(1988).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D00330; NOT_ANNOTATED_CDS.
 DR PIR: B28925; NKVLJ2.
 DR PFAM: PF00906; Hepatitis_core; 1.
 KW CORE PROTEIN; REPEAT.
 FT REPEAT 162 169
 FT REPEAT 170 177
 SQ SEQUENCE 183 AA; 21081 MW; 115BD9E3 CRC32;
 Query Match 44.98; Score 1336; DB 1; Length 183;
 Best Local Similarity 96.28; Pred. No. 5,22e-229;
 Matches 176; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Db 1 MDIDPKKFGASVELLSELPDFFPSYRDLDLTASALYREALSPCHSHHTALRAIL 60
 QY 1 MDIDPKKFGATVELLSELPDFFPSYRDLDLTASALYREALSPCHSHHTALRAIL 60
 Db 61 CMGELMNLATWGSNLEDPASRELIVSYVNVNMGKIRQLMLFHSICLTGRETVEIYLV 120
 QY 61 CMGELMNLATWGSNLEDPASRELIVSYVNVNMGKIRQLMLFHSICLTGRETVEIYLV 120
 Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPPRRRSQSPRRRSQRE 180
 QY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPPRRRSQSPRRRSQRE 180
 Db 181 SOC 183
 QY 181 SOC 183
 RESULT 5
 ID CORA.HPBV4 STANDARD; PRT; 183 AA.
 AC P03150; P03151.
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C.
 OS HEPATITIS B VIRUS (SUBTYPE ADP), HEPATITIS B VIRUS (SUBTYPE ADRA4),
 OS AND HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN INDONESIA/PIDW420).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ADR:
 RX MEDLINE: 83168919
 RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus
 RT DNA: subtype adr and adw."
 RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ADR4;
 RX MEDLINE: 83246570.
 RA FUJUYAMA A., MIYANOHARA A., NOZAKI C., YONEYAMA T., OHTOMO N.,
 RA MATSUBARA K.;
 RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
 RT adr."
 RL NUCLEIC ACIDS RES. 11:4601-4610(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ADM;
 RX MEDLINE: 89010694.
 RA OKAMOTO H., TSUDA F., SAKUGAMA H., SASTROSEWIGNO R.I., IMAI M.,
 RA MIYAKAWA Y., NAYUMI M.;

RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes."
 RL J. GEN. VIROL. 69:2575-2583(1988).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: V00867; NOT_ANNOTATED_CDS.
 DR EMBL: X01587; G59407; .
 DR EMBL: D00331; NOT_ANNOTATED_CDS.
 DR PIR: A93480; NKVLAS.
 DR PIR: B93460; NKVL44.
 DR PIR: C28925; NKVLJ3.
 DR PFAM: PF00906; Hepatitis_core; 1.
 KW CORE PROTEIN; REPEAT.
 FT REPEAT 162 169
 FT REPEAT 170 177
 SQ SEQUENCE 183 AA; 21095 MW; D2E185B9 CRC32;
 Query Match 44.88; Score 1334; DB 1; Length 183;
 Best Local Similarity 95.68; Pred. No. 1.30e-228;
 Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Db 1 MDIDPKKFGASVELLSELPDFFPSYRDLDLTASALYREALSPCHSHHTALRAIL 60
 QY 1 MDIDPKKFGATVELLSELPDFFPSYRDLDLTASALYREALSPCHSHHTALRAIL 60
 Db 61 CMGELMNLATWGSNLEDPASRELIVSYVNVNMGKIRQLMLFHSICLTGRETVEIYLV 120
 QY 61 CMGELMNLATWGSNLEDPASRELIVSYVNVNMGKIRQLMLFHSICLTGRETVEIYLV 120
 Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPPRRRSQSPRRRSQRE 180
 QY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPPRRRSQSPRRRSQRE 180
 Db 181 SOC 183
 QY 181 SOC 183
 RESULT 6
 ID CORA.HPBVJ STANDARD; PRT; 183 AA.
 AC P17351;
 DT 01-AUG-1990 (REL. 15, CREATED)
 DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C.
 OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN JAPAN/PJDM233).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 89010694.
 RA OKAMOTO H., TSUDA F., SAKUGAMA H., SASTROSEWIGNO R.I., IMAI M.,
 RA MIYAKAWA Y., NAYUMI M.;
 RT "Typing hepatitis B virus by homology in nucleotide sequence:
 RT comparison of surface antigen subtypes."
 RL J. GEN. VIROL. 69:2575-2583(1988).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D00329; NOT_ANNOTATED_CDS.
 DR PIR: A28925; NKVLJ1.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 21224 MW: F6B348B6 CRC32:

Query Match 44.4%; Score 1322; DB 1; Length 183;

Best Local Similarity 96.2%; Pred. No. 3.03e-226;
Matches 176; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROAIL 60
OY 1 MDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROAIL 60
Db 61 CGEGLMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEYLY 120
OY 61 CGEGLMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEYLY 120
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 180
OY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 180
Db 181 SOC 183
OY 181 SOC 183

RESULT 7 STANDARD: PRT; 185 AA.

ID CORA_HPBVT STANDARD: PRT; 185 AA.

AC P03149;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.

OS HEPATITIS B VIRUS (SUBTYPE ADW).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 83168919.

RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;

RT "The complete nucleotide sequences of the cloned hepatitis B virus

RL DNA; subtype adr and adv";

CC NUCLEIC ACIDS RES. 11:1747-1757(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC EMBL: V00866; NOT_ANNOTATED_CDS.

DR PIR: C93460; NKVLA6.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT

FT REPEAT 164 171

FT REPEAT 172 179

SQ SEQUENCE 185 AA: 21394 MW: 791E0381 CRC32:

Query Match 44.4%; Score 1321; DB 1; Length 185;

Best Local Similarity 96.2%; Pred. No. 4.78e-226;

Matches 178; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Db 1 MDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROAIL 60
OY 1 MDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROAIL 60
Db 61 CGEGLMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEYLY 120
OY 61 CGEGLMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEYLY 120

Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOS 180
OY 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOS 178
Db 181 RESOC 185
OY 179 RESOC 183

RESULT 8 STANDARD: PRT; 195 AA.

ID CORA_HPBVT STANDARD: PRT; 195 AA.

AC P29178;

DT 01-DEC-1992 (REL. 24, CREATED)

DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)

DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

OS HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 90169850.

RA BHAT R.A., ULRICH P.P., VYAS G.N.;

RT "Molecular characterization of a new variant of hepatitis B virus in

RL a persistently infected homosexual man.";

DR PIR: A37182; NKVLA3.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT

FT REPEAT 174 181

FT REPEAT 182 189

SQ SEQUENCE 195 AA: 22461 MW: AF3DB5F3 CRC32:

Query Match 44.1%; Score 1314; DB 1; Length 195;

Best Local Similarity 93.5%; Pred. No. 1.15e-224;

Matches 173; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 11 FGLDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROA 70
OY 212 LMSDIDPKKFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALROA 271
Db 71 LSCWELMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEY 130
OY 272 LSCWELMTLATWGNLLEDPASRDLYVNYVTNKGKIRQLMWFHISCLTFGRTVLEY 331
Db 131 LVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOS 190
OY 332 LVSFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOS 391
Db 191 RESOC 195
OY 392 RESOC 396

RESULT 9 STANDARD: PRT; 212 AA.

ID CORA_HPBVT STANDARD: PRT; 212 AA.

AC Q03495;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

OS HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE W4B).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 93346970.

RA NAUMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,

RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil

RL that expresses HBV surface antigen subtype adw4.";

J. GEN. VIROL. 74:1627-1632(1993).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: X69798; G59423; ..

DR PFAM: PF00906; Hepatitis_core: 1.

KW CORE PROTEIN: REPEAT.

FT DOMAIN 178 204 ARG-RICH.

FT REPEAT 184 203 3 X 5 AA REPEATS OF S-P-R-R-R.

FT REPEAT 184 188 1.

FT REPEAT 191 195 2.

FT REPEAT 199 203 3.

SO SEQUENCE 212 AA; 24234 MW; BCCDF263 CRC32;

Query Match 44.1%; Score 1313; DB 1; Length 212;

Best Local Similarity 92.9%; Pred. No. 1, 81e-224;

Matches 171; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

Db 29 GMDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 88

QY 213 SMDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 272

Db 89 LCMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 148

QY 273 LCMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 332

Db 149 VSGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSP 208

QY 333 VSGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSP 392

Db 209 ASOC 212

QY 393 ESOC 396

RESULT 10

ID CORA_HPBVL STANDARD; PRT: 183 AA.

AC P12901:

DT 01-OCT-1989 (REL. 12, CREATED)

DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)

DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

OS HEPATITIS B VIRUS (STRAIN LSH / CHIMPANZEE ISOLATE).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN 11

RP SEQUENCE FROM N.A.

RX MEDLINE: 88258473.

RA VAUDIN M., WOLSTEINHOE A.J., TSIOUVE K.N., ZUCKERMAN A.J.,

RA HARRISON T.J.:

RT "The complete nucleotide sequence of the genome of a hepatitis B

RT virus isolated from a naturally infected chimpanzee."

RL J. GEN. VIROL. 69:1383-1389(1988).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL: D00220; D1000603; ..

DR PIR: A28885; NKVLC.

DR PFAM: PF00906; Hepatitis_core: 1.

KW CORE PROTEIN: REPEAT.

FT REPEAT 162 169

FT REPEAT 170 177

SO SEQUENCE 183 AA; 20999 MW; IAF57C9 CRC32;

Query Match 43.9%; Score 1308; DB 1; Length 183;

Best Local Similarity 94.5%; Pred. No. 1, 76e-223;

Matches 173; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 60

QY 1 MDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 60

Db 61 CWMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 120

QY 61 CWMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 120

Db 121 SFGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSPA 180

QY 121 SFGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSPA 180

Db 181 SOC 183

QY 181 SOC 183

RESULT 11

ID CORA_HPBV2 STANDARD; PRT: 185 AA.

AC P03148:

DT 21-JUL-1986 (REL. 01, CREATED)

DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

OS HEPATITIS B VIRUS (SUBTYPE ADW2).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN 11

RP SEQUENCE FROM N.A.

RA VALENZUELA P., QUIROGA M., ZALDIVAR J., GRAY P., RUTTER W.J.;

RL (IN) FIELD B.N., JAKENISCH R., FOX C.F. (EDS.);

RL ANIMAL VIRUS GENETICS, PP. 57-70, ACADEMIC PRESS, NEW YORK (1980).

DR PIR: A94409; NKVLA3.

DR PFAM: PF00906; Hepatitis_core: 1.

KW CORE PROTEIN: REPEAT.

FT REPEAT 164 171

FT REPEAT 172 179

SO SEQUENCE 185 AA; 21304 MW; E2EA3360 CRC32;

Query Match 43.8%; Score 1305; DB 1; Length 185;

Best Local Similarity 95.7%; Pred. No. 6, 88e-223;

Matches 177; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Db 1 MDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 60

QY 1 MDIDPKFAGVAVLLSFLPSDFPSVRLDLDTSALYREALSPENCSPHNTALROAI 60

Db 61 CWMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 120

QY 61 CWMGELMTLASVGNLLEDPASRDLVSVYNTMGLKROLMLFHSICLTFGRVLEYL 120

Db 121 SFGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSP 180

QY 121 SFGWIRTTPAYRPPNAPILSTLPTETTVRRRGRSPRRRTSPRRRSQSPRRRSQSP 178

Db 181 RESOC 185

QY 179 RESOC 183

RESULT 12

ID CORA_HPBV9 STANDARD; PRT: 214 AA.

AC P17099:

DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

```

OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN 991).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RA SEQUENCE FROM N.A.
RX KOEHEL H.G., SCHUELER A., LOTTMANN S., THOMSEN R.;
RL SUBMITTED (FEB-1990) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51970; G60433; -
DR PIR: S10381; NKVLS.
DR PFAM: PF00906; Hepatitis_core; 1.
DR CORE PROTEIN: REPEAT.
FT REPEAT 193 200
FT REPEAT 201 208
SQ SEQUENCE 214 AA; 24722 MW; 786CD048 CRC32;

Query Match 43.8%; Score 1304; DB 1; Length 214;
Best Local Similarity 95.2%; Pred. No. 1,08e-222;
Matches 177; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

Db 29 GMDIDPKKEGATVETLSPDSEFPVRLDLPDASALYREALESPHCSPHHTALROAI 88
QY 213 SMDIDPKKEGATVETLSPDSEFPVRLDLPDASALYREALESPHCSPHHTALROAI 272
Db 89 LCMGELMTLATWGNLDEPDASRDLYVNYVNTNGLKIRQLMFRISYLTFFGRTVLEYL 148
QY 273 LCMGELMTLATWGNLDEPDASRDLYVNYVNTNGLKIRQLMFRISYLTFFGRTVLEYL 332
Db 149 SVSGVWIRTPAPYPNPAPILSTLPETTVARRRGASPRRRTPSPRRRSQSRRRRSQ 208
QY 333 SVSGVWIRTPAPYPNPAPILSTLPETTVARRR--GRSPRRTPSPRRRSQSRRRRSQ 390
Db 209 SRESOC 214
QY 391 SRESOC 396

RESULT 13
ID CORA_MHV1 STANDARD; PRT; 188 AA.
AC P03152;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 1 (MHV 1),
OS WOODCHUCK HEPATITIS VIRUS 7 (MHV 7),
OS WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE) (MHV 8),
OS WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE) (MHV 8),
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RA SEQUENCE FROM N.A.
RX SPECIES=WOODCHUCK HEPATITIS VIRUS 1;
RL MEDLINE: 82216969.
RA GALIBERT F., CHEN T.N., MANDART E.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT comparison with the hepatitis B virus sequence.";
RL J. VIROL. 41:51-65(1982).
RN [2]
RA SEQUENCE FROM N.A.
RX SPECIES=WOODCHUCK HEPATITIS VIRUS 7, AND WOODCHUCK HEPATITIS VIRUS 59;
RX MEDLINE: 88101359.
RA COHEN J.J., MILLER R.H., ROSENBLUM B., DENNISTON K., GERIN J.L.,
RA PORCELL R.H.;
RT "Sequence comparison of woodchuck hepatitis virus replicative forms
RT shows conservation of the genome.";

```

```

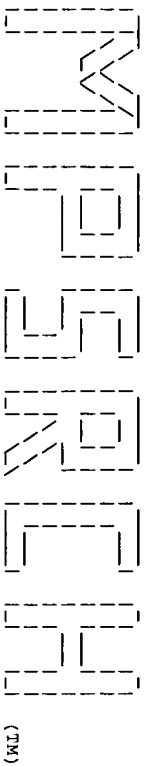
RL VIROLOGY 162:12-20(1988).
RN [3]
RA SEQUENCE FROM N.A.
RX SPECIES=WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE);
RX MEDLINE: 89184524.
RA GIROUX R., COTE P.J., HORNBUCKLE W.E., TENNANT B.C., GERIN J.L.,
RA PURCELL R.H., MILLER R.H.;
RT "Complete nucleotide sequence of a molecular clone of woodchuck
RT hepatitis virus that is infectious in the natural host.";
RL PROC. NATL. ACAD. SCI. U.S.A. 86:1846-1849(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02442; G336129; -
DR EMBL: M18752; G336140; -
DR EMBL: M19183; G336145; -
DR EMBL: J04514; G336149; -
DR PIR: A03713; NKVLC.
DR PIR: C32397; NKVLC.
DR PIR: C32397; NKVLC4.
DR PFAM: PF00906; Hepatitis_core; 1.
DR CORE PROTEIN: REPEAT.
FT REPEAT 167 174
FT REPEAT 175 182
SQ SEQUENCE 188 AA; 21693 MW; A667DB27 CRC32;

Query Match 31.5%; Score 938; DB 1; Length 188;
Best Local Similarity 67.0%; Pred. No. 6,28e-151;
Matches 126; Conservative 25; Mismatches 32; Indels 5; Gaps 1;

Db 1 MDIDPKKEGSSYQILNPLDFFPDLMALVDATALYEBELTGREHCSPHHTALROAI 60
QY 1 MDIDPKKEGATVETLSPDSEFPVRLDLPDASALYREALESPHCSPHHTALROAI 60
Db 61 CMDELTKLAMMSNTSEVRIIYNHVDITGLVROSLMHSCLTFGGTQVEFLY 120
QY 61 CMGELMTLATWGNLDEPDASRDLYVNYVNTNGLKIRQLMFRISYLTFFGRTVLEYL 120
Db 121 SVSGVWIRTPAPYPNPAPILSTLPETTVARRRGASPRRRTPSPRRRSQSRRRR 180
QY 121 SVSGVWIRTPAPYPNPAPILSTLPETTVARRR-----RSPRRTPSPRRRSQSRRRR 175
Db 181 SQSPSANC 188
QY 176 SQSRESOC 183

RESULT 14
ID CORA_HPV8 STANDARD; PRT; 217 AA.
AC P03153;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1990 (REL. 13, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS GROUND SQUIRREL HEPATITIS VIRUS (GSHV).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RA SEQUENCE FROM N.A.
RX MEDLINE: 84267998.
RA SEEGER C., GANEM D., YARMUS H.E.;
RT "Nucleotide sequence of an infectious molecularly cloned genome of
RT ground squirrel hepatitis virus.";
RL J. VIROL. 51:367-375(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 12:50:17 1999; Maspar time 26.61 Seconds
Tabular output not generated. 812.364 Million cell updates/sec

Title: >US38183-1-38183
Description: (1-396) from us38183-1-38183.pep
Perfect Score: 2977
Sequence: 1 MDIDPKKFGATVELSLFLP.....RRRSQSPRRRRSQSRESQC 396

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

sptrembl9
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 46.585; Variance 101.482; scale 0.459

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	46.2	212 14	089656	PRE-C/CORE.	2.16e-231
2	1374	46.2	212 14	088020	PRE-C/CORE.	3.37e-231
3	1374	46.2	212 14	089597	HBCAG.	3.37e-231
4	1373	46.1	183 14	089437	X, PREC AND C GENES (C	5.26e-231
5	1373	46.1	183 14	068008	X, PREC AND C GENES (F	5.26e-231
6	1373	46.1	212 14	067872	PRE C/C ORF.	5.26e-231
7	1372	46.1	212 14	011884	CORE ANTIGEN PRECURSOR	8.20e-231
8	1370	46.0	183 14	068056	X, PREC AND C GENES (L	1.99e-230
9	1370	46.0	212 14	068058	PRE-C/CORE.	1.99e-230
10	1370	46.0	212 14	068025	PRE-C/CORE.	1.99e-230
11	1369	46.0	212 14	067872	PRE C/C ORF.	3.11e-230
12	1369	46.0	212 14	067984	PRE-C/CORE.	3.11e-230
13	1369	46.0	212 14	068032	PRE-C/CORE.	3.11e-230
14	1368	46.0	212 14	068010	PRE-C/CORE.	4.85e-230
15	1366	45.9	212 14	068070	PRE-C/CORE.	1.18e-229
16	1366	45.9	212 14	068030	PRE-C/CORE.	1.18e-229
17	1364	45.8	183 14	067989	X, PREC AND C GENES (M	2.86e-229
18	1363	45.8	183 14	068048	X, PREC AND C GENES (F	4.46e-229
19	1363	45.8	212 14	068014	PRE-C/CORE.	4.46e-229
20	1363	45.8	212 14	068012	PRE-C/CORE.	4.46e-229

21	1363	45.8	212 14	067980	PRE-C/CORE.	4.46e-229
22	1362	45.8	212 14	068051	PRE-C/CORE.	6.96e-229
23	1360	45.7	212 14	089446	PRE-C/CORE.	1.69e-228
24	1360	45.7	212 14	068016	PRE-C/CORE.	1.69e-228
25	1360	45.7	212 14	096845	COMPLETE GENOME.	1.69e-228
26	1359	45.6	212 14	068053	PRE-C/CORE.	2.64e-228
27	1357	45.6	212 14	068075	PRE-C/CORE.	6.41e-228
28	1357	45.6	212 14	068077	PRE-C/CORE.	6.41e-228
29	1355	45.5	212 14	068045	PRE-C/CORE.	1.56e-227
30	1352	45.4	183 14	067931	CORE PROTEIN.	5.90e-227
31	1351	45.4	183 14	067946	C ANTIGEN.	9.20e-227
32	1351	45.4	212 14	089719	PRE-C/CORE.	9.20e-227
33	1349	45.3	183 14	067997	X, PREC AND C GENES (B	2.24e-226
34	1348	45.3	183 14	068037	X, PREC AND C GENES (D	3.49e-226
35	1348	45.3	183 14	068060	X, PREC AND C GENES (M	1.32e-225
36	1345	45.2	183 14	067973	X, PREC AND C GENES (T	5.43e-226
37	1347	45.2	212 14	067912	PRE-CORE/CORE PROTEIN.	5.43e-226
38	1343	45.1	183 14	068023	X, PREC AND C GENES (C	3.12e-225
39	1342	45.1	183 14	067964	X, PREC AND C GENES (F	5.06e-225
40	1344	45.1	212 14	092918	PRECORE PROTEIN PRECUR	2.06e-225
41	1344	45.1	212 14	068035	PRE-C/CORE.	2.06e-225
42	1341	45.0	183 14	068064	X, PREC AND C GENES (L	7.80e-225
43	1341	45.0	183 14	068003	X, PREC AND C GENES (S	7.80e-225
44	1341	45.0	212 14	068042	PRE-C/CORE.	7.80e-225
45	1341	45.0	212 14	081115	HBCAG,HBCAG.	7.80e-225

ALIGNMENTS

RESULT	ID	1	PRELIMINARY:	PRT:	212 AA.
DC	089656;				
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)				
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)				
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)				
DE	PRE-C/CORE.				
GN	PRE-C/CORE.				
OS	HEPATITIS B VIRUS.				
OC	VERUSSES; RETROID VIRUSSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRATIN-PATIENT LICHERI-1'85.				
RL	LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;				
RL	SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRATIN-AYW4.				
RA	PLUCIENNICZAK A.;				
RL	SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 81012091.				
RA	GALLIERI F., MANDART E., FITOUSSI F., TIOLLAIS P., CHARNAY P.;				
RT	"Nucleotide sequence of the hepatitis B virus genome (subtype ayw)				
RT	cloned in E. coli."				
RL	NATURE 281:646-650(1979).				
RA	[4]				
RP	SEQUENCE FROM N.A.				
RA	BORISOVA G.P., GREN E.Y., TSIBINGIN V.V., KURAIN R.A.;				
RA	DISHLER A.V., GREN E.Y., BYCHKO V.V., PUSHKO P.M., KALIS Y.V.;				
RA	DOKL. BIOCHEM. 279:386-390(1985).				
DR	EMBL: X85290; G736116; -.				
DR	EMBL: X85300; G736152; -.				
DR	EMBL: X85313; G736196; -.				
DR	EMBL: X85283; G736090; -.				
DR	EMBL: X85306; G736174; -.				
DR	EMBL: X85316; G527437; -.				
DR	EMBL: J02203; G329642; -.				
DR	EMBL: X85312; G736193; -.				
DR	PFAM: PF00906; Hepatitis_core; 1.				
KW	STGNAL.				
SO	SEQUENCE	212 AA;	24350 MM;	71EA2C82 CRC32;	

Query Match 46.2%; Score 1375; DB 14; Length 212;
Best Local Similarity 99.5%; Pred. No. 2.16e-231;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 88
:|||||
QY 213 SMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 272
:|||||
DB 89 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 148
:|||||
QY 273 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 332
:|||||
DB 149 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
:|||||
QY 333 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 392
:|||||
DB 209 ESQC 212
:|||||
QY 393 ESQC 396

RESULT 2
ID 068020 PRELIMINARY: PRT: 212 AA.
AC 068020;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA LAI M.E., MAZOLENTI A.P., PORU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85281; G736119; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 46.2%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3.37e-231;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 88
:|||||
QY 213 SMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 272
:|||||
DB 89 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 148
:|||||
QY 273 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 332
:|||||
DB 149 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
:|||||
QY 333 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 392
:|||||
DB 209 ESQC 212
:|||||
QY 393 ESQC 396

RESULT 3
ID 089597 PRELIMINARY: PRT: 212 AA.
AC 089597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HBCAG.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SUB-TYPE AYW;

Query Match 46.2%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3.37e-231;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 88
:|||||
QY 213 SMDIDPKFEGATVLLSFLPSDFPSVRLDITASALYREALSPHCSPHHTALROAI 272
:|||||
DB 89 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 148
:|||||
QY 273 LCWGLMTLATWGVNLEDPASRDLYSVNTNMGKLFROLMFHISCLTFGRETYEYL 332
:|||||
DB 149 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
:|||||
QY 333 VSFQWIRTPPAPRPAPNPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 392
:|||||
DB 209 ESQC 212
:|||||
QY 393 ESQC 396

RESULT 4
ID 089437 PRELIMINARY: PRT: 183 AA.
AC 089437;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (CASTA 2).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA LAI M.E., MAZOLENTI A.P., PORU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85289; G736113; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 46.1%; Score 1373; DB 14; Length 183;

Query Match 46.1%; Score 1373; DB 14; Length 183;

Accession	Score	DB	Length	Indels	Gaps
RL EMBL: X65258; G59436; -	46.1%;	Score 1372;	DB 14;	Length 212;	
DR PFAM: PF00906; Hepatitis_core: 1.	Best Local Similarity 98.9%;	Pred. No. 5,26e-231;			
SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;	Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;				
Db 29 GMDIDPKYKFGATVLLSLFSDFFPSVRDLDTASALYREALESPHCSPHNTALROAI 88					
QY 213 SMDIDPKYKFGATVLLSLFSDFFPSVRDLDTASALYREALESPHCSPHNTALROAI 272					
Db 89 LCMGELMTLATWGVNLEDPASRLVSYNTNGLKFLRQLMFHISCLTGGRTVTEYL 148					
QY 273 LCMGELMTLATWGVNLEDPASRLVSYNTNGLKFLRQLMFHISCLTGGRTVTEYL 332					
Db 149 VAFGWIPTPPAYPPNAPILSTLPETVYRRGRSPRRRTPPSPRRRSQSPRRRSQSR 208					
QY 333 VSPFGWITPPAYPPNAPILSTLPETVYRRGRSPRRRTPPSPRRRSQSPRRRSQSR 392					
Db 209 ESQC 212					
QY 393 ESQC 396					
RESULT 7 PRELIMINARY; PRT; 212 AA.					
ID 011884					
AC 011884					
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)					
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)					
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)					
DE CORE ANTIGEN PRECURSOR.					
OS HEPATITIS B VIRUS.					
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.					
OC [1]					
RP SEQUENCE FROM N. A.					
RA RAO B.S., CASEY J.L., RINAUDO J.S., KORBA B.E.;					
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.					
DR EMBL: U95511; G2182120; -					
DR PFAM: PF00906; Hepatitis_core: 1.					
SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;					
Query Match 46.1%; Score 1372; DB 14; Length 212;					
Best Local Similarity 98.9%; Pred. No. 8,20e-231;					
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;					
Db 29 GMDIDPKYKFGATVLLSLFSDFFPSVRDLDTASALYREALESPHCSPHNTALROAI 88					
QY 213 SMDIDPKYKFGATVLLSLFSDFFPSVRDLDTASALYREALESPHCSPHNTALROAI 272					
Db 89 LCMGELMTLATWGVNLEDPASRLVSYNTNGLKFLRQLMFHISCLTGGRTVTEYL 148					
QY 273 LCMGELMTLATWGVNLEDPASRLVSYNTNGLKFLRQLMFHISCLTGGRTVTEYL 332					
Db 149 VAFGWIPTPPAYPPNAPILSTLPETVYRRGRSPRRRTPPSPRRRSQSPRRRSQSR 208					
QY 333 VSPFGWITPPAYPPNAPILSTLPETVYRRGRSPRRRTPPSPRRRSQSPRRRSQSR 392					
Db 209 EPQC 212					
QY 393 ESQC 396					
RESULT 8 PRELIMINARY; PRT; 183 AA.					
ID 068066					
AC 068066					
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)					
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)					
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)					
DE X, PREC AND C GENES (LICHERI 2).					
OS CORE.					
OS HEPATITIS B VIRUS.					

```
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2/87;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 6F38AA3B CRC32;

Query Match
Best Local Similarity 98.4%; Score 1370; DB 14; Length 183;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
QY 1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
Db 61 CAGDMLTATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 120
QY 61 CAGDMLTATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 120
Db 121 SEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 180
QY 121 SEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 180
Db 181 SOC 183
QY 181 SOC 183

RESULT 9
ID 068068 PRELIMINARY; PRT: 212 AA.
AC 068068;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3/90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
QY 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
Db 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
QY 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
Db 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
QY 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
Db 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
QY 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
QY 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
Db 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
QY 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
Db 209 ESQC 212
QY 209 ESQC 212
Db 393 ESQC 396
QY 393 ESQC 396

RESULT 10
ID 068025 PRELIMINARY; PRT: 212 AA.
AC 068025;
```

```
AC 068025;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG/3;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
QY 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
Db 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
QY 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
Db 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
QY 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
Db 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
QY 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
QY 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
Db 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
QY 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
Db 209 ESQC 212
QY 209 ESQC 212
Db 393 ESQC 396
QY 393 ESQC 396

RESULT 11
ID 067872 PRELIMINARY; PRT: 212 AA.
AC 067872;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C, HBSAG POSITIVE, HBV SUBTYPE ATW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24368 MW; A6BD9DA1 CRC32;

Query Match
Best Local Similarity 98.4%; Score 1369; DB 14; Length 212;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
QY 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
Db 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
QY 213 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 272
Db 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
QY 89 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 148
Db 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
QY 273 LCGELMTLATWGVNLEDPASRDLVSVYNTNMGKLFROLIMFHSCLTFGRTVIEYL 332
Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
QY 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 208
Db 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
QY 333 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRPSRRRSQSPRRRSQSR 392
Db 209 ESQC 212
QY 209 ESQC 212
```

QY 393 ESOC 396

RESULT 12
ID 067984 PRELIMINARY: PRT: 212 AA.
AC 067984;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT VITTORINA'92;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85256; G736052; -
DR PFAM: PF00906; Hepatitis_core: 1.
SO SEQUENCE 212 AA; 24322 MW; 83D9780B CRC32;

Query Match
Best Local Similarity 98.9%; Score 1369; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
QY :|||||
213 SMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
273 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 332

Db 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY :|||||
333 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY :|||||
393 ESOC 396

RESULT 13
ID 068032 PRELIMINARY: PRT: 212 AA.
AC 068032;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CHICHINE-2'86;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85296; G736139; -
DR PFAM: PF00906; Hepatitis_core: 1.
SO SEQUENCE 212 AA; 24308 MW; 69D87B53 CRC32;

Query Match
Best Local Similarity 98.4%; Score 1369; DB 14; Length 212;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
QY :|||||
213 SMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
273 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 332

Db 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY :|||||
333 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY :|||||
393 ESOC 396

QY 273 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 332
Db 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY :|||||
333 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY :|||||
393 ESOC 396

RESULT 14
ID 068010 PRELIMINARY: PRT: 212 AA.
AC 068010;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100; -
DR PFAM: PF00906; Hepatitis_core: 1.
SO SEQUENCE 212 AA; 24292 MW; 04A4D12D CRC32;

Query Match
Best Local Similarity 98.4%; Score 1368; DB 14; Length 212;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
QY :|||||
213 SMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 272

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
273 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFHISCLTFGRETVIEYL 332

Db 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY :|||||
333 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 392

Db 209 ESOC 212
QY :|||||
393 ESOC 396

RESULT 15
ID 068070 PRELIMINARY: PRT: 212 AA.
AC 068070;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GORDO/84;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210; -
DR PFAM: PF00906; Hepatitis_core: 1.
SO SEQUENCE 212 AA; 24376 MW; 80F52D0F CRC32;

Query Match
Best Local Similarity 98.9%; Score 1366; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEGATVELLSFLPSDFPFSVRDLDTASALYREALESPHCSPHHTALROAI 88
:|||||
QY 213 SMDIDPYKEGATVELLSFLPSDFPFSVRDLDTASALYREALESPHCSPHHTALROAI 272
Db 89 LCWGEIMTLATWGVNLEDDPASRDLYVSYVNTNMGKLFROLMFHISCLTFGRETVEYL 148
:|||||
QY 273 LCWGEIMTLATWGVNLEDDPASRDLYVSYVNTNMGKLFROLMFHISCLTFGRETVEYL 332
Db 149 VSEGVWIRTPPAYRPPNAPILLLPETTVVRRGRSPRRRTPSPPRRRSQSPRRRSQSR 208
:|||||
QY 333 VSEGVWIRTPPAYRPPNAPILLLPETTVVRRGRSPRRRTPSPPRRRSQSPRRRSQSR 392
Db 209 ESQC 212
:|||||
QY 393 ESQC 396

Search completed: Thu Dec 16 12:52:36 1999
Job time : 139 secs.

US3835-1-38183
MDIDPKKEFGATVELLSFLPSDFPPSVRDLLDTASEIITRDGELLQMDFFPEHLVDLQSLMDIDP
YKEFGATVELLSFLPSDFPPSVRDLLDTASEIITRDGELLQMDFFPEHLVDLQSLMDIDP
LEDPASRDLVSVYNTNMGLKFRQLMFIHISCLTFGRETVEYLVSGVWIRTPPAYRPPNAPILSTLPE
TTVRRRGSRPRRTTSPRRRSQSPRRRSQSRRESQCI

US 3835-1-38183
seq 38 residue 1-35
seq 38 residue 1-183
seq 1-183 not 30 translation

 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 12:57:17 1999; MasPar time 11.85 Seconds

Tabular output not generated. 445,188 Million cell updates/sec

Title: >US3835-1-38183
 Description: (1-248) from us3835-1-38183.pep
 Perfect Score: 1854
 Sequence: 1 MDIDPKKEGATVELLSFLP.....RRRSQSPRRRSQSNESOC 248

Scoring table: PAM 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.490; Variance 173.849; scale 0.193

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1375	74.2	194	32	W50242	Hepatitis B virus pre
2	1375	74.2	212	32	W50250	Hepatitis B virus pre
3	1375	74.2	346	5	R27473	S12/core protein.
4	1363	73.3	184	1	P80959	Hepatitis B virus sub
5	1359	73.3	193	32	W50241	Hepatitis B virus pre
6	1358	73.2	183	32	W50251	Hepatitis B virus p21
7	1353	73.0	183	20	W09044	Hepatitis B virus cor
8	1349	72.8	183	1	R05635	Hepatitis B antigen.
9	1349	72.8	183	5	P00041	Sequence of core anti
10	1344	72.5	397	20	W09048	Plasmid pHBV DN AA en
11	1338	72.2	184	5	P00004	Sequence of core anti
12	1336	72.1	196	8	R40806	Hepatitis B core / pV
13	1337	72.1	208	8	R40808	Hepatitis B core / pV
14	1334	72.0	183	4	P40311	Hepatitis B virus core
15	1334	72.0	183	8	R40805	Hepatitis B core prot
16	1334	72.0	183	13	R68868	Hepatitis B virus pol

17	1334	72.0	183	12	R62869	Hepatitis B virus cor	2.26e-96
18	1333	71.9	198	8	R40807	Hepatitis B core / IL	2.73e-96
19	1328	71.6	293	10	R55286	Presl full length cor	7.00e-96
20	1325	71.5	183	13	R68866	Hepatitis B virus cor	1.23e-95
21	1325	71.5	214	1	P90702	Deduced amino acid se	1.23e-95
22	1325	71.5	289	20	W09049	Plasmid pHBV DN BB en	1.23e-95
23	1324	71.4	185	6	R30861	Hepatitis B core anti	1.49e-95
24	1324	71.4	185	10	R55284	Deduced sequence of f	1.49e-95
25	1324	71.4	185	6	R30780	Native HBcAg protein.	1.49e-95
26	1324	71.4	185	6	R31025	Native HBcAg protein.	1.49e-95
27	1321	71.3	185	26	P31061	Core antigen of HBV.	2.52e-95
28	1320	71.2	183	13	R68869	Hepatitis B virus pol	3.16e-95
29	1317	71.0	185	6	R30781	Altered HBcAg protein	5.57e-95
30	1317	71.0	185	6	R31026	Altered HBcAg protein	5.57e-95
31	1314	70.9	183	13	R68867	Hepatitis B virus pol	9.81e-95
32	1313	70.8	183	13	R68870	Hepatitis B virus pol	1.18e-94
33	1311	70.7	185	6	R30784	Altered HBcAg protein	1.73e-94
34	1311	70.7	185	6	R31029	Altered HBcAg protein	1.73e-94
35	1311	70.7	185	6	R30865	Modified hepatitis B	1.73e-94
36	1309	70.6	185	6	R30862	Hepatitis B core anti	2.52e-94
37	1308	70.6	185	6	R31028	Altered HBcAg protein	3.04e-94
38	1308	70.6	185	6	R30783	Altered HBcAg protein	3.04e-94
39	1308	70.6	185	6	R30864	Modified hepatitis B	3.04e-94
40	1303	70.3	214	1	P80961	HBV core antigen enco	7.80e-94
41	1302	70.2	203	1	P82872	HBcAg/Beta-Gal fusion	9.42e-94
42	1287	69.4	198	32	W50252	Hepatitis B virus E a	1.59e-92
43	1263	68.1	183	17	R98878	Hepatitis B virus E a	1.46e-90
44	1259	67.9	183	17	R98883	Hepatitis B virus E a	3.11e-90
45	1258	67.9	183	17	R98884	Hepatitis B virus E a	3.75e-90

ALIGNMENTS

RESULT 1
 ID W50242 standard; Protein; 194 AA.
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p22 polypeptide Met-p22.
 KW Viral replication; inhibitor: HBV; nucleocapsid; gene therapy;
 KM hepatocyte; liver; Met-p22.
 OS Hepatitis B virus.
 OS Synthetic.
 FH Key
 FT Protein
 FT 2..194
 FT /label= p22
 PN W09809649-A1.
 PD 12-MAR-1998.
 PE 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (CEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI; 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 11: Page 40; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein with an added N-terminal Met residue. p22 is produced by
 CC elimination of the 19-amino acid leader peptide from the 25 kDa
 CC full-length HBV precore protein (see W50250). Evidence is provided
 CC that HBV replication is inhibited in the presence of high levels of
 CC HBV precore or precore-related proteins. These proteins can be
 CC incorporated into HBV nucleocapsids along with the p21 core protein
 CC (see W50251), which is the usual nucleocapsid component, and
 CC thereby render the nucleocapsids deficient in encapsidating HBV
 CC pregenomic RNA. Thus, over-expression of the precore proteins, or
 CC certain variants of them, leads to transdominant inhibition of HBV
 CC replication. Suitable inhibitory proteins include p25 (see W50250),
 CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
 CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously

to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 194 AA;

Query Match 74.2%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 9,896-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmdldpykxkfegatvellsfipdsfipsvrdldtasalyrealsepshcphthalrgai 70
 :|||||
 QY 65 SMDIDPKKEGATVELLSFIPDSFIPSVRDLDLTSALYREALSEPSCPHHTALRQAI 124
 Db 71 lcwgelmtlatwgvnledpasrdlsvsyvntnmglkfrqlwfhiscitfgretvieyl 130
 :|||||
 QY 125 LCWGEIIMTLATWGVNLEDPASRDLSVSYVNTNMGKFRQLMFHISCITFGRETVEYL 184
 Db 131 vsfgvwlrtppayrppnapilsltpetvrrrgsprrrtsprrrrsgsrrrrsgsr 190
 :|||||
 QY 185 VSGVGVWIRTPPAYRPPNAPILSLTPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 244
 Db 191 esgc 194
 :|||
 QY 245 ESQC 248

RESULT 2
 ID W50250 standard; Protein; 212 AA.
 AC W50250:

DE 28-SEP-1998 (first entry)
 KW Hepatitis B virus precore p25 polypeptide.
 KW Viral replication; inhibitor: HBV; nucleocapsid; gene therapy;
 KW hepatocyte; liver; p25 protein.
 OS Hepatitis B virus.
 PN M09809649-A1.
 PD 12-MAR-1998.
 PF 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PT Claim 15; Page 35; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SQ Sequence 212 AA;

Query Match 74.2%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 9,896-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmdldpykxkfegatvellsfipdsfipsvrdldtasalyrealsepshcphthalrgai 88
 :|||||
 65 SMDIDPKKEGATVELLSFIPDSFIPSVRDLDLTSALYREALSEPSCPHHTALRQAI 124

Db 89 lcwgelmtlatwgvnledpasrdlsvsyvntnmglkfrqlwfhiscitfgretvieyl 148
 :|||||
 QY 125 LCWGEIIMTLATWGVNLEDPASRDLSVSYVNTNMGKFRQLMFHISCITFGRETVEYL 184
 Db 149 vsfgvwlrtppayrppnapilsltpetvrrrgsprrrtsprrrrsgsrrrrsgsr 208
 :|||||
 QY 185 VSGVGVWIRTPPAYRPPNAPILSLTPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 244
 Db 209 esgc 212
 :|||
 QY 245 ESQC 248

RESULT 3
 ID R27473 standard; Protein; 346 AA.
 AC R27473:

DE 24-FEB-1993 (first entry)
 DE S12/core protein.
 KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KW vaccinia virus; I3L; promoter; NVVAC; recombinant; HBV L;
 KW large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
 KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KW deletion loci; recipient loci.
 OS Synthetic.
 FH key
 FT region
 FT 1.108
 FT /label S1
 FT region
 FT 109..163
 FT /label S2
 FT region
 FT 164..346
 FT /label Core
 PN W09215672-A.
 PD 17-SEP-1992.
 PF 09-MAR-1992; U01906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRO-) VIROGENETICS CORP.
 PA Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Limbach KJ, Norton EK, Paoletti E, Perkins ME, Pincus SE,
 PI Riviere M, Tarantola J, Taylor J;
 DR WPI: 92-331718/40.
 DR N-PSDB; Q29105.
 PT vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure; Fig 13; 45pp; English.
 CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the 13L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NVVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, lpsAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NVVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NVVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC Q35501-864.
 SQ Sequence 346 AA;

Query Match 74.2%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 9,896-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 nmldidykxkfegatvellsfipdsfipsvrdldtasalyrealsepshcphthalrgai 222


```

OY      :|||||
65  SMDIDPKKEGATVELLSFSDFFPSVRLDPTASALYREALSEPHCSPHHTALRQAI 124
Db      223  lcgelmtlatwgvnlgedpasrdlvsvyvdtnmgkfrglwfhmscltfgrvteyle1 282
OY      125  lcgelmtlatwgvnlgedpasrdlvsvyvdtnmgkfrglwfhmscltfgrvteyle1 184
Db      283  vsfgvwlrtppayrppnapllscipetlvrrgrspprrtsprrrsqsprrrrsqsqr 342
OY      185  vsfgvwlrtppayrppnapllscipetlvrrgrspprrtsprrrsqsprrrrsqsqr 244
Db      343  esgc 346
OY      245  ESQC 248

RESULT 4
ID      P80959 standard: protein: 184 AA.
AC      P80959;
DE      19-NOV-1990 (first entry)
DT      Hepatitis B virus subtype ayw. core protein.
KW      conjugate: fusion protein; hepatitis B core protein; HBV subtype ayw.;
KM      T cell stimulating polypeptide; vaccines.
OS      Synthetic.
PN      EP-271302-A.
PD      15-JUN-1988.
PF      07-DEC-1987: 310725.
PR      07-OCT-1987: US-106538.
RA      07-OCT-1987: US-939617.
PA      (SCR-) Scripts Clinic Res.
PI      Thornton GB, Moriarty AM, Millich DR, McLachlan A.;
DR      WPI: 88-163287/24.
PT      New conjugates and fusion proteins of immunogenic polypeptide -
PT      and hepatitis B core antigen and T cell stimulating polypeptide
PT      corep. to core antigen, useful in vaccines.
PS      Disclosure: P: English.
CC      This sequence contains the T cell stimulating epitopes, amino acid
CC      residue 1-55 and 70-140. It is believed that the regions 1-44 and
CC      70-140 do not contain determinants that suppress T cell activation.
CC      Polypeptides essentially consisting of 15 to 55 amino acids
CC      corresponding to the above mentioned HBV regions are T cell
CC      stimulating. Coupling a polypeptide immunogen to such sequences,
CC      e.g. by using a bifunctional reagent which forms a disulphide link,
CC      improves its immunogenicity. These are useful in vaccines and can be used
CC      therapeutically to improve T cell response to HBcAg in infected
CC      subjects.
CC      See also P80896-P80898 and P80951-P80959.
SO      Sequence 184 AA.

Query Match      73.5%: Score 1363; DB 1; Length 184;
Best Local Similarity 99.5%: Pred. No. 9.52e-99;
Matches 183; Conservative 0; Mismatches 0; Indels 1; Gaps 1:

```

```

DT      28-SEP-1998 (first entry)
DE      Hepatitis B virus precore p22 polypeptide.
KW      viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
KM      hepatocyte; liver; p22 protein.
OS      Hepatitis B virus.
FH      Key      Location/Qualifiers
FT      Region      82..98
FT      W09809649-A1.
FT      P12-MAR-1998.
PD      03-SEP-1997: 015500.
PR      03-SEP-1996: US-025370.
PA      (GEO) GEN HOSPITAL CORP.
PI      Melegari M, Scaglioni PP, Wands JR.
DR      WPI: 98-193325/17.
PT      DNA encoding proteins which can be incorporated with wild type
PT      nucleocapsid subunit(s) into a viral nucleocapsid - useful for
PT      inhibition of viral replication, especially hepatitis B virus
PS      Claim 9; Page 34-35; 60pp: English.
CC      This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
CC      protein that is produced by elimination of the 19-amino acid leader
CC      peptide from the 25 kDa full-length HBV precore protein (see
CC      W50250). Evidence is provided that HBV replication is inhibited in
CC      the presence of high levels of HBV precore or precore-related
CC      proteins. These proteins can be incorporated into HBV nucleocapsids
CC      along with the p21 core protein (see W50251), which is the usual
CC      nucleocapsid component, and thereby render the nucleocapsids
CC      deficient in encapsidating HBV pregenomic RNA. Thus, over-
CC      expression of the precore proteins, or certain variants of them,
CC      leads to transdominant inhibition of HBV replication. Suitable
CC      inhibitory proteins include p25 (see W50250), p22, Met-p22 (see
CC      W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
CC      (see W50238). Heterologous peptides (see W50244-49) may be
CC      inserted into the p22 and Met-p22 polypeptides. The inhibitory
CC      proteins can be produced by recombinant methods using claimed
CC      expression vectors and host cells. They can be provided exogenously
CC      to the target cells for use in inhibiting HBV replication.
CC      Alternatively, a nucleic acid construct that directs overexpression
CC      of an inhibitory protein in target cells is used for the gene
CC      therapy of HBV infection.
SO      Sequence 193 AA.

Query Match      73.3%: Score 1359; DB 32; Length 193;
Best Local Similarity 98.4%: Pred. No. 2.02e-98;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0:

```

FT /note= "immunodominant region"
 PN WO9809649-A1.
 PD 12-MAR-1998.
 PR 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (GEMO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 98-193325/17.
 PR DNA encoding proteins which can be incorporated with wild type
 PR nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PR inhibition of viral replication, especially hepatitis B virus
 PS Disclosure: Page 41-42: 60pp: English.
 CC This is the 21 kDa (p21) protein of hepatitis B virus (HBV) that
 CC assembled into a 180 kDa subunit nucleocapsid structure that
 CC promotes viral replication. Evidence is provided that HBV
 CC replication is inhibited in the presence of high levels of HBV
 CC precore or precore-related proteins. These proteins can be
 CC incorporated into HBV nucleocapsids along with the p21 core protein
 CC and thereby render the nucleocapsids deficient in encapsidating HBV
 CC pregenomic RNA. Thus, over-expression of the precore proteins or
 CC certain variants of them, leads to transdominant inhibition of HBV
 CC replication. Suitable inhibitory proteins include p25 (see
 CC W50250), p22 (see W50241), Met-p22 (see W50242), p18 (see W50235),
 CC Met-p18 (see W50237) and Met-p18-Het (see W50238). These proteins
 CC can be produced by recombinant methods using claimed expression
 CC vectors and host cells, and can be provided exogenously to target
 CC cells for use in inhibiting HBV replication. Alternatively, a
 CC nucleic acid construct that directs overexpression of an inhibitory
 CC protein in target cells is used for the gene therapy of HBV
 CC infection.
 SO Sequence 183 AA:

Query Match 73.2%; Score 1358; DB 32; Length 183;
 Best Local Similarity 98.9%; Pred. No. 2,44e-98;
 Matches 181; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 1 mdidpykefgatvelslsfpsdfsvrdlldtasalyrealesphesphhtalrgail 60
 QY 66 MDIDPYKEFGATVELSLSPDSDFSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 125
 DB 61 cweelmtlatwgvnllepdasrdlvsvyvdntmgjlkfrqilwfhscldfgetvleylv 120
 QY 126 CWGELMTLATWGVNLLEPDASRDLVSVYVNTNMGKLFQRLMFHISCLFGEETVLEYLV 185
 DB 121 sfqvwlrtpayrppnapilslpetvrrrgsprrrrtsprrrrsgsrrrrsgsre 180
 QY 186 SFGVWLRTPAYRPPNAPILSLPETVRRRGSPRRRRTPSPRRRRSOSPPRRRSQSRE 245
 DB 181 sgq 183
 QY 246 SOC 248
 OS Hepatitis B virus.
 KW core protein: replication; antiviral; gene therapy.
 FH Key
 FT Location/Qualifiers
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 71-180 in mutant
 FT polypeptides of the invention (Claim 23)"
 FT misc_difference 81..180
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 81 and 180 in mutant
 FT polypeptides of the invention (Claim 5)"
 FT misc_difference 171..180
 FT /note= "C-terminus of core protein is at any amino

FT acid position between 171 and 180 in mutant
 FT polypeptides of the invention (Claim 6)"
 FT misc_difference 174..180
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 174 and 180 in mutant
 FT polypeptides of the invention (Claim 26)"
 FT misc_difference 178
 FT /note= "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT misc_difference 178
 FT /note= "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT misc_difference 172..183
 FT /note= "amino acid residues 172-183, pref. 174-180,
 FT are deleted from the core protein in mutant
 FT polypeptides of the invention"

PN WO9700698-A1.
 PD 09-JAN-1997.
 PR 20-JUN-1996; U10602.
 PR 20-JUN-1995; US-017814.
 PA (GEMO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 97-087176/08.
 DR N-PSDB: 149594.
 PR New method for inhibiting the replication of hepadnaviruses -
 PR comprises introducing a mutant polypeptide with a mutated core
 PR protein or corresponding nucleic acid, for treating, e.g. hepatitis
 PR B

PS Claim 5; Page 55-56; 83pp; English.
 CC Alteration of the C-terminus of the core protein (W09044) of
 CC hepatitis B virus (HBV) or other hepadnavirus creates a mutant
 CC polypeptide capable of reducing replication of the wild-type virus
 CC by a dominant negative mechanism. The inhibitory effect is
 CC species-specific and is achieved by deletion of a few C-terminal
 CC amino acids from the core protein, and/or by joining the core
 CC protein to a hepadnavirus surface protein (see also W09045),
 CC creating a core-surface fusion. Such mutant polypeptides (see
 CC also W09046-50) can be used to treat hepadnavirus, e.g. HBV,
 CC infection.
 SO Sequence 183 AA:

Query Match 73.0%; Score 1353; DB 20; Length 183;
 Best Local Similarity 98.9%; Pred. No. 6,28e-98;
 Matches 181; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 1 mdidpykefgatvelslsfpsdfsvrdlldtasalyrealesphesphhtalrgail 60
 QY 66 MDIDPYKEFGATVELSLSPDSDFSVRDLLDTASALYREALESPEHCSPHHTALRQAIL 125
 DB 61 cweelmtlatwgvnllepdasrdlvsvyvdntmgjlkfrqilwfhscldfgetvleylv 120
 QY 126 CWGELMTLATWGVNLLEPDASRDLVSVYVNTNMGKLFQRLMFHISCLFGEETVLEYLV 185
 DB 121 sfqvwlrtpayrppnapilslpetvrrrgsprrrrtsprrrrsgsrrrrsgsre 180
 QY 186 SFGVWLRTPAYRPPNAPILSLPETVRRRGSPRRRRTPSPRRRRSOSPPRRRSQSRE 245
 DB 181 sgq 183
 QY 246 SOC 248

RESULT 8
 ID R05635 standard: protein; 183 AA.
 AC R05635.
 DT 30-OCT-1990 (first entry)
 DE Hepatitis B antigen.
 KW Hepatitis B virus; vaccine; HBV; ds.
 OS Synthetic.
 PN EP-374869-A.
 PD 27-JUN-1990.

```
PF 1-JAN-1989: 123526.
PR 22-DEC-1978: GB-049907.
PR 27-DEC-1978: GB-050039.
PR 01-NOV-1979: GB-037910.
PA (BIOJ) Biogen Inc.
PI Murray K, Schaller HE;
DR WPI: 90-195067/26.
DR N-PSDB: Q04799.
PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
PT used in detection of infection and in vaccine prodn.
PS Disclosure: 4pp: English.
CC Polypeptide fragments encoded by the DNA sequence are antigenic for
CC HBV and may be used as a vaccine or in detection. Peptides may be
CC cultured in a suitable bacterial host such as E.coli.
CC Fragments of the sequence are also claimed as being antigenically
CC useful.
SQ Sequence 183 AA:

Query Match 72.8%; Score 1349; DB 1; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.33e-97;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdidykefgatvellstfspdffpsvrdlldtaalyrdalesepshphltargall 60
    |||
QY 66 MDIDYKEFGATVELLSPLSPDFPSVRLDLDTASALYREALSEPHSPHHTALRQAIL 125
    |||
Db 61 cwgdlmtlatwgvnlledpasrdlvsvyntnmglkfqrllwfhscitfgretyleylv 120
    |||
QY 126 CWGELMTLATWGVNLLEDPASRDLVSVYNTNMGKRFQRLMFHISCLTFGRETYIEIYLV 185
    |||
Db 121 sfywlrtppayrpnaplsltpetlvrrrgsprrrrrpsrrrrsgsrrrrsqr 180
    |||
QY 186 SFGWIRTPPAYRPNPADILSTLPETVVRRGSPRRRTPSPRRRSQSPRRRSQSRE 245
    |||
Db 181 sqc 183
    |||
QY 246 SQC 248

RESULT 9
ID P00041 standard; Protein: 183 AA.
AC P00041:
DT 14-OCT-1992 (first entry)
DE Sequence of core antigen.
KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
OS Hepatitis B virus.
PN EP--13828-A.
PD 06-AUG-1980.
PF 21-DEC-1979: 303017.
PR 22-DEC-1978: GB-049907.
PR 27-DEC-1978: GB-050039.
PR 01-NOV-1979: GB-037910.
PA (BIOJ ) BIOGEN NV.
PI Murray K, Schaller HE;
DR WPI: 80-57268C/33.
DR N-PSDB: N00003.
PT Recombinant DNA coding for polypeptide - have specificity of
PT hepatitis B viral antigens in detection or antibody stimulation
PS Example: Figs 3-4: 43pp: English.
CC Human serum from a single HBsAg positive, HBeAg positive donor
CC (serotype adyM) was used to prep. a DNA-contg. pellet which was
CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
CC The labelled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prep'd. by the processes are deposited at the
CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg:
CC HBV-Kpn I dc: Tetr Amps HBV+.
SQ Sequence 183 AA:

Query Match 72.8%; Score 1349; DB 5; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.33e-97;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 1 mdidykefgatvellstfspdffpsvrdlldtaalyrdalesepshphltargall 60
    |||
QY 66 MDIDYKEFGATVELLSPLSPDFPSVRLDLDTASALYREALSEPHSPHHTALRQAIL 125
    |||
Db 61 cwgdlmtlatwgvnlledpasrdlvsvyntnmglkfqrllwfhscitfgretyleylv 120
    |||
QY 126 CWGELMTLATWGVNLLEDPASRDLVSVYNTNMGKRFQRLMFHISCLTFGRETYIEIYLV 185
    |||
Db 121 sfywlrtppayrpnaplsltpetlvrrrgsprrrrrpsrrrrsgsrrrrsqr 180
    |||
QY 186 SFGWIRTPPAYRPNPADILSTLPETVVRRGSPRRRTPSPRRRSQSPRRRSQSRE 245
    |||
Db 181 sqc 183
    |||
QY 246 SQC 248

RESULT 10
ID W09048 standard; Protein: 397 AA.
AC W09048:
DT 11-APR-1997 (first entry)
DE Plasmid pHBV DN AA encoded HBV dominant negative polypeptide.
KW Hepadnavirus; woodchuck hepatitis virus; hepatitis delta virus;
KW HBV; ground squirrel hepatitis B virus; duck hepatitis B virus;
KW core protein; replication; antiviral; gene therapy; pHBV DN AA.
OS Hepatitis B virus.
FH Key Location/Qualifiers
FT region 1..179
FT /label= HBV_core
FT /note= "positions 1-179 correspond to amino acids
FT 1-179 of HBV core protein"
FT region 180..397
FT /note= "positions 180-397 correspond to amino
FT acids 9-226 of HBV surface protein"
FN W09700698-A1.
PD 09-JAN-1997.
PF 20-JUN-1996; U10602.
PR 20-JUN-1995; US-017814.
PA (GENO ) GEN HOSPITAL CORP.
PI Melegari M, Scaglioni PP, Wands JR;
DR WPI: 97-087176/08.
DR N-PSDB: T49598.
PT New method for inhibiting the replication of hepadnaviruses -
PT comprises introducing a mutant polypeptide with a mutated core
PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
PT B
PS Disclosure: Page 46-48: 83pp: English.
CC Plasmid pHBV DN AA carries an insert (T49598) coding for a protein
CC (W09048) comprising a hepatitis B virus (HBV) core protein (see
CC also W09044) fused in-frame at amino acid 179 with the HBV surface
CC protein (see also W09045) at amino acid 8. Plasmid pHBV DN BB
CC (T49599) expresses the HBV core fused at amino acid 175 to the
CC surface protein at amino acid 112 (W09049). pHBV DN AA was at
CC least as potent an inhibitor of HBV replication as construct
CC pHBV DN (T49597). pHBV DN BB was less inhibitory than pHBV DN.
CC Vectors expressing hepadnavirus dominant negative core mutants can
CC be utilized in the gene therapy of viral infections.
SQ Sequence 397 AA:

Query Match 72.5%; Score 1344; DB 20; Length 397;
Best Local Similarity 100.0%; Pred. No. 3.43e-97;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdidykefgatvellstfspdffpsvrdlldtaalyrdalesepshphltargall 60
    |||
QY 66 MDIDYKEFGATVELLSPLSPDFPSVRLDLDTASALYREALSEPHSPHHTALRQAIL 125
    |||
Db 61 cwgdlmtlatwgvnlledpasrdlvsvyntnmglkfqrllwfhscitfgretyleylv 120
    |||
QY 126 CWGELMTLATWGVNLLEDPASRDLVSVYNTNMGKRFQRLMFHISCLTFGRETYIEIYLV 185
    |||
Db 121 sfywlrtppayrpnaplsltpetlvrrrgsprrrrrpsrrrrsgsrrrrsqr 179
```

BA. (NISW) NISSHIN OIL MILLS LTD.

```
QY 64 L$MDIDPKFEGATVELLSFLPSDFP$VRDLDLTASALREALESPENCSPHHTALROA 12
```

```

OY 124 ILKWEELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMHISCLTGRETVLEY 183
DB 144 LVSFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 203
OY 184 LVSFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 243
DB 204 resqc 208
OY 244 RESQC 248

```

RESULT 14

ID P40311 standard; Protein; 183 AA.

```

AC P40311;
DT 18-AUG-1992 (first entry)
DE Hepatitis virus core antigen.
KW HBcAg; vaccine; diagnosis; HBV infection.
OS Hepatitis B virus.
PN J59074985-A.
PD 27-APR-1984.
PF 19-OCT-1982; 183432.
PR 19-OCT-1982; JP-183432.
PA (TAKE ) TAKEDA CHEMICAL IND KK.
DR MPI: 84-143231/23.
PT DNA used in prevention of infections by hepatitis virus B -
  comprises structural gene of hepatitis virus adr B surface antigen
  coding gene and at least 1 virus core antigen structural gene.
PS Disclosure: Fig 3; 13pp; Japanese.
CC The sequence is that of hepatitis virus core antigen (HBcAg). It
  can be used as a vaccine for the prevention of infections by
  hepatitis B virus (HBV) and also in the diagnosis of early stages
  of HBV infection. See also P40310.
SQ Sequence 183 AA;

```

Query Match 72.0%; Score 1334; DB 4; Length 183;

Best Local Similarity 95.6%; Pred. No. 2,266-96;

Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

DB 1 mdldpykefgasvvelsfpsdfffpsirdldtasalyrealespeshcphthalrgall 60
OY 66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESPESHCPHHTALRQAIL 125
DB 61 CWGELMNLATWGVNLEDPASRDVSVYNTNMGKFRQLMHISCLTGRETVLEYLV 120
OY 126 CWGELMNLATWGVNLEDPASRDVSVYNTNMGKFRQLMHISCLTGRETVLEYLV 185
DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 180
OY 186 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 245
DB 181 sqc 183
OY 246 SOC 248

```

RESULT 15

ID R40805 standard; Protein; 183 AA.

```

AC R40805;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core protein.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Hepatitis B virus.
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISM ) NISSHIN OIL MILLS LTD.
DR MPI: 93-277479/35.
DR N-PSDB: Q47735.
PT Recombinant plasmid for high immunogenicity virus - contains
  recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
  virus and exotic genes

```

```

PS Disclosure: Fig 6; 12pp; Japanese.
CC The hepatitis B core gene is recombinated with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.
SQ Sequence 183 AA;

```

Query Match 72.0%; Score 1334; DB 8; Length 183;

Best Local Similarity 95.6%; Pred. No. 2,266-96;

Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

DB 1 mdldpykefgasvvelsfpsdfffpsirdldtasalyrealespeshcphthalrgall 60
OY 66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALESPESHCPHHTALRQAIL 125
DB 61 CWGELMNLATWGVNLEDPASRDVSVYNTNMGKFRQLMHISCLTGRETVLEYLV 120
OY 126 CWGELMNLATWGVNLEDPASRDVSVYNTNMGKFRQLMHISCLTGRETVLEYLV 185
DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 180
OY 186 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRPSRRRSQSRRRSQS 245
DB 181 sqc 183
OY 246 SOC 248

```

Search completed: Thu Dec 16 13:00:09 1999

Job time : 172 secs.

 NWSETH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Dec 16 13:00:27 1999; Maspar time 12.64 Seconds
 Tabular output not generated. 786.145 Million cell updates/sec

Title: >US3835-1-38183
 Description: (1-248) from us3835-1-38183.pep
 Perfect Score: 1854
 Sequence: 1 MDIDYKEFGATVELLSFLP.....RRRRSQSPRRRRSQSRNSQC 248

Scoring table: PAM 150
 Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: plr60
 1:plr1 2:plr2 3:plr3 4:plr4

Statistics: Mean 47.523; Variance 119.912; scale 0.396

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	74.2	212	1	NKVLAH e antigen precursor / 4.43e-199	
2	1373	74.1	183	2	S53207 core antigen - hepati 9.53e-199	
3	1374	74.1	212	2	S53221 e antigen precursor / 6.50e-199	
4	1374	74.1	212	2	S532204 e antigen precursor / 6.50e-199	
5	1373	74.1	212	2	S53270 e antigen precursor / 9.53e-199	
6	1370	73.9	183	2	S53270 core antigen - hepati 3.01e-198	
7	1370	73.9	212	2	S53216 e antigen precursor / 3.01e-198	
8	1370	73.9	212	2	S53272 e antigen precursor / 3.01e-198	
9	1369	73.8	212	2	S53163 e antigen precursor / 4.42e-198	
10	1369	73.8	212	2	S520746 e antigen precursor / 4.42e-198	
11	1369	73.8	212	2	S53225 e antigen precursor / 4.42e-198	
12	1368	73.8	212	2	S53198 e antigen precursor / 6.48e-198	
13	1366	73.7	212	2	S53223 e antigen precursor / 1.39e-197	
14	1366	73.7	212	2	S53274 e antigen precursor / 1.39e-197	
15	1364	73.6	183	2	S53169 core antigen - hepati 3.00e-197	
16	1363	73.5	212	2	S53247 core antigen - hepati 4.40e-197	
17	1363	73.5	212	2	S53202 e antigen precursor / 4.40e-197	
18	1363	73.5	212	2	S53159 e antigen precursor / 4.40e-197	
19	1363	73.5	212	2	S53200 e antigen precursor / 4.40e-197	
20	1362	73.5	212	2	S53251 e antigen precursor / 6.46e-197	
21	1360	73.4	212	2	S53204 e antigen precursor / 1.39e-196	
22	1360	73.4	212	2	S53227 e antigen precursor / 1.39e-196	
23	1359	73.3	212	2	S53253 e antigen precursor / 2.04e-196	

24	1358	73.2	211	1	NKVLAI e antigen precursor / 2.99e-196
25	1357	73.2	212	2	S53281 e antigen precursor / 4.39e-196
26	1355	73.1	212	1	NKVLBH e antigen precursor / 9.44e-196
27	1355	73.1	212	2	S53242 e antigen precursor / 9.44e-196
28	1351	72.9	183	2	S53129 core antigen - hepati 4.37e-195
29	1351	72.9	212	2	S53240 e antigen precursor / 4.37e-195
30	1349	72.8	183	1	NKVLAI2 core antigen - hepati 9.41e-195
31	1349	72.8	183	2	S53181 core antigen - hepati 9.41e-195
32	1348	72.7	183	2	S53232 core antigen - hepati 1.38e-194
33	1348	72.7	183	2	S53260 core antigen - hepati 1.38e-194
34	1345	72.5	183	2	S53152 core antigen - hepati 1.36e-194
35	1344	72.5	212	2	S53229 e antigen precursor / 6.39e-194
36	1343	72.4	183	2	S53214 core antigen - hepati 9.37e-194
37	1342	72.4	183	2	S53140 core antigen - hepati 1.37e-193
38	1341	72.3	183	2	S53267 core antigen - hepati 2.02e-193
39	1341	72.3	183	2	S53189 core antigen - hepati 2.02e-193
40	1341	72.3	212	2	S53228 e antigen precursor / 2.02e-193
41	1340	72.3	212	2	S53257 e antigen precursor / 2.96e-193
42	1339	72.2	183	2	S53137 core antigen - hepati 4.34e-193
43	1339	72.2	212	2	S52651 e antigen precursor / 4.34e-193
44	1338	72.2	212	2	S53236 e antigen precursor / 6.36e-193
45	1337	72.1	212	1	NKVLJ2 e antigen precursor / 9.33e-193

ALIGNMENTS

RESULT 1
 ENTRY 1
 TITLE NKVLAH #type complete
 e antigen precursor / core antigen - hepatitis B virus
 (subtype ayw4, isolate hb321 and others)
 ALTERNATE_NAMES HBe antigen precursor / HBeC antigen; pre-C/C antigen
 CONTAINS core antigen; e antigen
 ORGANISM #formal_name hepatitis B virus; HBV
 #variety subtype ayw4, isolate hb321; isolate patient Ferracuti '83;
 - isolate patient Castag '83; isolate patient Sanna '84;
 - isolate patient Licheri '1'85; isolate patient Flore '1'86;
 isolate patient Licheri '83
 18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change 10-Oct-1997

DATE 18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change 10-Oct-1997
 ACCESSIONS S47405; S53191; S53209; S53234; S53264; S53249; S53262;
 S53277; A03711

REFERENCE S47404

#authors Plucienicak, A.

#submission submitted to the EMBL Data Library, August 1994

#description Molecular cloning and sequencing of two complete genomes of
 Polish isolates of human hepatitis B virus.

#accession S47405

#molecule_type DNA

#residues 1-212 #label PLU

#cross-references EMBL:X5316; NID:9527435; PID:9527437

#experimental_source subtype ayw4, isolate hb321

REFERENCE S53112

#authors Lal, M.E.; Marzoleni, A.P.; Portu, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53191

#molecule_type DNA

#residues 1-212 #label LA1

#cross-references EMBL:X85283; NID:9736088; PID:9736090

#experimental_source isolate patient Ferracuti '83

#accession S53209

#molecule_type DNA

#residues 1-212 #label LA2

#cross-references EMBL:X85290; NID:9736114; PID:9736116

#experimental_source isolate patient Castag '83

#accession S53234

#molecule_type DNA

#residues 1-212 #label LA3

#cross-references EMBL:X85300; NID:9736150; PID:9736152

#experimental_source isolate patient Sanna '84

#accession S53264

#molecule_type DNA

#residues 1-212 #label LA4

#cross-references EMBL:X85313; NID:9736194; PID:9736196

##experimental_source isolate patient Licheri-1'85
#accession S53249
##molecule_type DNA
##residues 1-212 ##label LA5
##cross-references EMBL:X85306; NID:q736172; PID:q736174
##experimental_source isolate patient Fiore-1'86
#accession S53262
##molecule_type DNA
##residues 1-212 ##label LA6
##cross-references EMBL:X85312; NID:q736191; PID:q736193
##experimental_source isolate patient Licheri'83
#accession S53277
##molecule_type DNA
##residues 30-212 ##label LA7
##cross-references EMBL:X85317; NID:q736211; PID:q736214
##experimental_source patient Giordo-2'86
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be processed

REFERENCE A93214
#authors Galibert, F.; Mandart, E.; Flitoussi, F.; Tlollais, P.;
#journal Nature (1979) 281:646-650
#title Nucleotide sequence of the hepatitis B virus genome (subtype
ayw) in E. coli.
#cross-references M01D:81012091
#accession A03711
##molecule_type DNA
##residues 1-212 ##label GAL
##cross-references GB:j02203; NID:q329640; PID:q329642
##experimental_source subtype ayw

GENETICS C
#gene #superfamily hepatitis B virus core antigen
CLASSIFICATION alternative initiators: core protein
FEATURE
KEYWORDS
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link ENG #status
predicted #label ECP
SUMMARY #length 212 #molecular-weight 24350 #checksum 782

Query Match 74.2%; Score 1373; DB 1; Length 212;
Best Local Similarity 99.5%; Pred. No. 4,43e-199;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKRGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHHTALROAIL 88
:|||||
QY 65 SMDIDPKRGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHHTALROAIL 124
:|||||

Db 89 LCGELMTLATWGVNEDPASRDVLSYVNTNMGKLFROLMFHISCLTFGRETVEYL 148
:|||||
QY 125 LCGELMTLATWGVNEDPASRDVLSYVNTNMGKLFROLMFHISCLTFGRETVEYL 184
:|||||

Db 149 VSGVWIRTPPAVPAPPLSTLPETTVVRRGRSPRRRSPRRRSOSR 208
:|||||
QY 185 VSGVWIRTPPAVPAPPLSTLPETTVVRRGRSPRRRSPRRRSOSR 244
:|||||

Db 209 ESOC 212
:|||||
QY 245 ESOC 248

RESULT 2
ENTRY S53207 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
Casta-2'87)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castaa-2'87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997 0.

ACCESSIONS S53207
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53207
##molecule_type DNA
##residues 1-183 ##label LA1
##cross-references EMBL:X85289; NID:q736110; PID:q736113
##experimental_source isolate patient Castaa-2'87
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS C
#gene #superfamily hepatitis B virus core antigen
CLASSIFICATION core protein
KEYWORDS #length 183 #molecular-weight 21102 #checksum 2165
SUMMARY

Query Match 74.1%; Score 1373; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 9,53e-199;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKRGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHHTALROAIL 60
:|||||
QY 66 MDIDPKRGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHHTALROAIL 125
:|||||

Db 61 CWGELMTLATWGVNEDPASRDVLSYVNTNMGKLFROLMFHISCLTFGRETVEYL 120
:|||||
QY 126 CWGELMTLATWGVNEDPASRDVLSYVNTNMGKLFROLMFHISCLTFGRETVEYL 185
:|||||

Db 121 SFGVWIRTPPAVPAPPLSTLPETTVVRRGRSPRRRSPRRRSOSR 180
:|||||
QY 186 SFGVWIRTPPAVPAPPLSTLPETTVVRRGRSPRRRSPRRRSOSR 245
:|||||

Db 181 SOC 183
:|||||
QY 246 SOC 248

RESULT 3
ENTRY S53211 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Castag-1'85 and others)
ALTERNATE_NAMES HBC antigen; Hbe antigen precursor / HBC antigen; pre-C/C
antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Castag-1'85; isolate patient Ferracuti-1'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997
ACCESSIONS S53211; S53197
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53211
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85291; NID:q736117; PID:q736119
##experimental_source isolate patient Castag-1'85
#accession S53197
##molecule_type DNA
##residues 30-212 ##label LAW
##cross-references EMBL:X85284; NID:q736095; PID:q736098
##experimental_source isolate patient Ferracuti-1'89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS C
#gene #superfamily hepatitis B virus core antigen
CLASSIFICATION alternative initiators: core protein
FEATURE
KEYWORDS #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG


```

179-212      #domain carboxy1-terminal propeptide #link EAG #status
              predicted #label ECP
SUMMARY      #length 212 #molecular-weight 24336 #checksum 746

Query Match  74.1%: Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%: Pred. No. 6,50e-199;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSLPSPFFPSVSRDLDTASALREALSEPHCSPHHTALROAI 88
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 SMDIDPKKEGATVELLSLPSPFFPSVSRDLDTASALREALSEPHCSPHHTALROAI 124
LCMGDLTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 148
125 LCMGELMTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 184
149 VSFSGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSOSR 208
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
185 VSFSGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSOSR 244
Db 209 ESOC 212
|||||
QY 245 ESOC 248

RESULT 4
ENTRY S32204 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1000)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#subtype ayw, patient C1000
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997

ACCESSIONS S32204
REFERENCE S32202
AUTHORS Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.;
Gerok, W.; Rasenack, J.
#submission submitted to the EMBL Data Library, March 1993
#description Identification and sequence analysis of hepatitis B virus DNA
in immunological negative infection.
#accession S32204
##molecule_type DNA
##residues 1-212 ##label PRE
##cross-references EMBL:X172702; NID:9286927; PID:9286930
##experimental_source subtype ayw, patient C1000
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label CAG\
179-212 #domain carboxy1-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular-weight 24336 #checksum 770

Query Match 74.1%: Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%: Pred. No. 6,50e-199;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSLPSPFFPSVSRDLDTASALREALSEPHCSPHHTALROAI 88
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 SMDIDPKKEGATVELLSLPSPFFPSVSRDLDTASALREALSEPHCSPHHTALROAI 124
LCMGELMTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 148
125 LCMGELMTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 184
89 LCMGELMTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 148
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 LCMGELMTLTATWGVNLDEPARDLVVSYNTMGAKFQOLLMPFHSCITFERREVIETL 184

```

Db	149	VSFGWIRTPAYPPNAPNPISTLPETTVVRRGRSRRRTPPSPRRRSQSPRRRSQSR	208
Oy	185	VSFGWIRTPAYPPNAPNPISTLPETTVVRRGRSRRRTPPSPRRRSQSPRRRSQSR	244
Db	209	ESQC 212	
Oy	245	ESQC 248	
RESULT	5		
ENTRY		S20750	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus (subtype ayw, patient CI)	
ALTERNATE_NAMES		HBe antigen precursor / HBe antigen; pre-C/C antigen	
CONTAINS		core antigen; e antigen	
ORGANISM		#formal_name hepatitis B virus, HBV	
#variety		subtype ayw, patient CI	
DATE		20-Sep-1995	#sequence_revision 20-Feb-1995
		08-Sep-1997	#text_change
ACCESSIONS		S20750	
REFERENCE		S20745	
authors		Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.	
#submission		submitted to the EMBL Data Library, March 1992	
#description		Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic liver disease.	
#accession		S20750	
#molecule_type		DNA	
#residues		1-212	#label LAI
#cross-references		EMBL:X65558; NID:g59434; PID:g5436	
#experimental_source		subtype ayw, patient CI	
GENETICS			
CLASSIFICATION		C	
FEATURE		#superfamily hepatitis B virus core antigen	
KEYWORDS		alternative initiators; core protein	
1-29			
30-212		#domain signal sequence #status predicted #label SIG\	
30-178		#product core antigen #status predicted #label CAG\	
179-212		#product e antigen #status predicted #label ENG\	
		#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP	
SUMMARY		#length 212 #molecular_weight 24348 #checksum 147	
Query Match		74.1%; Score 1373; DB 2; Length 212;	
Best Local Similarity		98.9%; Pred. No. 9,53e-199;	
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;			
Db	29	GMDIDPKYKCAIYELLISLPDSFFSVRLDLTASALYEALESPPHCSPHHTALROAI	88
Oy	65	SMDDIPKFEFAIYELLISLPDSFFSVRLDLTASALYEALESPPHCSPHHTALROAI	124
Db	89	LCMGELMTLAWGVNVEDPASRLVYVSYNTNGLKFRQLLMFHISCLTFGRETVLEYL	148
Oy	125	LCMGELMTLAWGVNVEDPASRLVYVSYNTNGLKFRQLLMFHISCLTFGRETVLEYL	184
Db	149	VAFGWIRTPAYPPNAPNPISTLPETTVVRRGRSRRRTPPSPRRRSQSPRRRSQSR	208
Oy	185	VSFGWIRTPAYPPNAPNPISTLPETTVVRRGRSRRRTPPSPRRRSQSPRRRSQSR	244
Db	209	ESQC 212	
Oy	245	ESQC 248	
RESULT	6		
ENTRY		S53270	#type complete
TITLE		core antigen - hepatitis B virus (isolate patient Licheri-2'87)	
ALTERNATE_NAMES		HBC antigen	
CONTAINS		core antigen	
ORGANISM		#formal_name hepatitis B virus, HBV	
#variety		isolate patient Licheri-2'87	
DATE		08-Jul-1995	#sequence_revision 03-Aug-1995
			#text_change

ACCESSIONS	08-Sep-1997
REFERENCE	S53270
#authors	Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission	Submitted to the EMBL Data Library, March 1995
#accession	S53270
##molecule-type	DNA
##residues	1-183 ##label LAI
##cross-references	EMBL:X85314; NID:g736201; PID:g736204
##experimental_source	isolate patient Licheri-2/87
##note	due to a stop codon between the alternative initiators the e antigen precursor cannot be produced
GENETICS	
CLASSIFICATION	C
KEYWORDS	#superfamily hepatitis B virus core antigen core protein
SUMMARY	length 183 #molecular-weight 21102 #checksum 2199
Query Match	73.9%; Score 1370; DB 2; Length 183;
Best Local Similarity	98.4%; Pred. NO. 3.01e-198;
Matches	180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db	1 MDIDPKFGATVELSLPDSFPYRDLDITRASALYRDALESPHCSPHTALRAIL 60 66 MDIDPKFGATVELSLPDSFPYRDLDITRASALYRDALESPHCSPHTALRAIL 125
Db	61 CWMGLMTLATWGVNLDPASRDVLVSVYNTNMGKFRQLLMFHSICLTGRETVEIYL 120
Db	126 CWMGLMTLATWGVNLDPASRDVLVSVYNTNMGKFRQLLMFHSICLTGRETVEIYL 185
Db	121 SFGWIRTPPAYRPPNAPILSTLETVYVRRKGTTPRRTPSPRRRSQSPRRRSQRE 180
Qy	186 SFGWIRTPPAYRPPNAPILSTLETVYVRRKGTTPRRTPSPRRRSQSPRRRSQRE 245
Db	181 SOC 183
Qy	246 SOC 248
RESULT	7
ENTRY	S53216 #type complete
TITLE	e antigen precursor / core antigen - hepatitis B virus (isolate patient Castag'3)
ALTERNATE_NAMES	HBe antigen precursor / HBeC antigen; pre-C/C antigen
CONTAINS	core antigen; e antigen
ORGANISM	#formal_name hepatitis B virus; HBV
#variety	isolate patient Castag'3
DATE	08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997
ACCESSIONS	S53216
REFERENCE	S53112
#authors	Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission	submitted to the EMBL Data Library, March 1995
#accession	S53216
##molecule-type	DNA
##residues	1-212 ##label LAI
##cross-references	EMBL:X85293; NID:g736124; PID:g736126
##experimental_source	isolate patient Castag'3
GENETICS	
CLASSIFICATION	C
KEYWORDS	#superfamily hepatitis B virus core antigen alternative initiators; core protein
FEATURE	
1-29	#domain signal sequence #status predicted #label SIG\
30-212	#product core antigen #status predicted #label CAG\
30-178	#product e antigen #status predicted #label EAG\
179-212	#domain carboxyl-terminal propeptide #link EXG #status predicted #label ECP
SUMMARY	#length 212 #molecular-weight 24363 #checksum 752
Query Match	73.9%; Score 1370; DB 2; Length 212;
Best Local Similarity	98.9%; Pred. NO. 3.01e-198;
Matches	182; Conservative 1; Indels 0; Gaps 0;

Db	29	GMDIDPYKEFGATVELLSFLPSDFPSPYRDLLDNASALYREALESPHCSPHNTALROAI	88
Qy	65	SMDDIDPYKEFGATVELLSFLPSDFPSPYRDLLDNASALYREALESPHCSPHNTALROAI	124
Db	89	LCWGLMLTLAWGVNLEDPAASRDLYSVYVNTNGGLFROLIMPHISCLTFGRRTVEYL	148
Qy	125	LCWGLMLTLAWGVNLEDPAASRDLYSVYVNTNGGLFROLIMPHISCLTFGRRTVEYL	184
Db	149	VSEFGWIRTPAYRPNPAPILSTLPETTVVARRGRSPRRRTPSRRRRSSQSR	208
Qy	185	VSEFGWIRTPAYRPNPAPILSTLPETTVVARRGRSPRRRTPSRRRRSSQSR	244
Db	209	ESQC 212	
Qy	245	ESQC 248	
RESULT	8		
ENTRY	553272	#type complete	
TITLE	e antigen precursor / core antigen - hepatitis B virus		
ALTERNATE_NAMES	(isolate patient Licheri-3'90)		
CONTAINS	HBE antigen precursor / HBC antigen; pre-C/C antigen		
ORGANISM	core antigen; e antigen		
#variety	#formal_name hepatitis B virus, HBV		
DATE	isolate patient Licheri-3'90		
ACCESSIONS	08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997		
REFERENCE	S53272		
#authors	Lai, M.E.; Mazzei, A.P.; Porru, A.; Balestrieri, A.		
#submission	submitted to the EMBL data library, March 1995		
#accession	S53272		
#molecule_type	DNA		
#residues	1-212 #label LAI		
GENETICS	#cross-references EMBL:X85315; NID:q736205; PID:q736207		
	#experimental_source isolate patient Licheri-3'90		
CLASSIFICATION	C		
KEYWORDS	#superfamily hepatitis B virus core antigen		
FEATURE	alternative initiators; core protein		
1-29			
30-212	#domain signal sequence #status predicted #label SIG\		
30-178	#product e antigen #status predicted #label CAG\		
179-212	#domain carboxyl-terminal propeptide #link ENG #status predicted #label ECP		
SUMMARY	#length 212 #molecular_weight 24366 #checksum 446		
Query Match	73.9%; Score 1370; DB 2; Length 212;		
Best Local Similarity	98.9%; Pred. No.3.0e-198;		
Matches 182; Conservative	1; Mismatches 1; Indels 0; Gaps 0		
Db	29	GMDIDPYKEFGATVELLSFLPSDFPSPYRDLLDNASALYREALESPHCSPHNTALROAI	88
Qy	65	SMDDIDPYKEFGATVELLSFLPSDFPSPYRDLLDNASALYREALESPHCSPHNTALROAI	124
Db	89	LCWGLMLTLAWGVNLEDPAASRDLYSVYVNTNGGLFROLIMPHISCLTFGRRTVEYL	148
Qy	125	LCWGLMLTLAWGVNLEDPAASRDLYSVYVNTNGGLFROLIMPHISCLTFGRRTVEYL	184
Db	149	VSEFGWIRTPAYRPNPAPILSTLPETTVVARRGRSPRRRTPSRRRRSSQSR	208
Qy	185	VSEFGWIRTPAYRPNPAPILSTLPETTVVARRGRSPRRRTPSRRRRSSQSR	244
Db	209	ESQC 212	
Qy	245	ESQC 248	
RESULT	9		
ENTRY	553163	#type complete	
TITLE	e antigen precursor / core antigen - hepatitis B virus		

ALTERNATE_NAMES (isolate patient Vittorina'92)
CONTAINS HBE antigen precursor / HBC antigen; pre-C/C antigen
ORGANISM core antigen; e antigen
#formal_name hepatitis B virus, HBV
#variety isolate patient Vittorina'92
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997
ACCESSIONS S53163
REFERENCE S53112
#authors Lal, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53163
#molecule_type DNA
#residues 1-212 ##label LAI
##cross-references EMBL:X85256; NID:g736050; PID:g736052
##experimental_source isolate patient Vittorina'92
GENETICS
CLASSIFICATION C
KEYWORDS #superfamily hepatitis B virus core antigen
#alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular_weight 24322 #checksum 842
Query Match 73.8%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 4,42e-198;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 29 GMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 88
:|||||
65 SMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 124
:|||||
Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 148
:|||||
125 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 184
:|||||
Db 149 VSEGVWTRTPAPYRPPAPILSTLPETTVVRRGRSPRRRSPRRRSOSR 208
:|||||
185 VSEGVWTRTPAPYRPPAPILSTLPETTVVRRGRSPRRRSPRRRSOSR 244
:|||||
Qy 209 ESOC 212
:|||||
Qy 245 ESOC 248
RESULT 10
ENTRY S20746 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C)
ALTERNATE_NAMES HBE antigen precursor / HBC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient C
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997
ACCESSIONS S20746
REFERENCE S20745
#authors Lal, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.;
Porru, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with
HBsAg negative chronic liver disease.
#accession S20746
#molecule_type DNA
#residues 1-212 ##label LAI
##cross-references EMBL:X65257; NID:g959429; PID:g959431
##experimental_source subtype ayw, patient C
GENETICS
#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS #alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular_weight 24348 #checksum 506
Query Match 73.8%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4,42e-198;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 29 GMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 88
:|||||
65 SMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 124
:|||||
Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 148
:|||||
125 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 184
:|||||
Db 149 VSEGVWTRTPAPYRPPAPILSTLPETTVVRRGRSPRRRSPRRRSOSR 208
:|||||
185 VSEGVWTRTPAPYRPPAPILSTLPETTVVRRGRSPRRRSPRRRSOSR 244
:|||||
Db 209 ESOC 212
:|||||
Qy 245 ESOC 248
RESULT 11
ENTRY S53225 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Chighine-2'86)
ALTERNATE_NAMES HBE antigen precursor / HBC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Chighine-2'86
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997
ACCESSIONS S53225
REFERENCE S53112
#authors Lal, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53225
#molecule_type DNA
#residues 1-212 ##label LAI
##cross-references EMBL:X85296; NID:g736137; PID:g736139
##experimental_source isolate patient Chighine-2'86
GENETICS
CLASSIFICATION C
KEYWORDS #superfamily hepatitis B virus core antigen
#alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular_weight 24308 #checksum 655
Query Match 73.8%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4,42e-198;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 29 GMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 88
:|||||
65 SMDIDPKFEGATVELLSFSPDFPSVRDLDTASALYREALSPHCSPHNTALROAI 124
:|||||
Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 148
:|||||
125 LCWGLMTLATWGVNLEDPASRDVSVYNTNGLKFKRLMLFHSICLTFGRTVIEYL 184
:|||||

```

Db      149 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
      |||
QY      185 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 244
      |||
Db      209 ESOC 212
      |||
QY      245 ESOC 248

RESULT  12
ENTRY   S53198
TITLE   #type complete
        e antigen precursor / core antigen - hepatitis B virus
        (isolate patient Ferracuti-2'90)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORANISM #formal_name hepatitis B virus, HBV
        isolate patient Ferracuti-2'90
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
        08-Sep-1997

ACCESSIONS S53198
REFERENCE S53112
AUTHORS Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
SUBMISSION submitted to the EMBL Data Library, March 1995
#accession S53198
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85285; NID:q736099; PID:q736100
#experimental_source isolate patient Ferracuti-2'90

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
        predicted #label ECP
SUMMARY #length 212 #molecular-weight 24292 #checksum 593

Query Match 73.8%; Score 1368; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 6,48e-198;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db      29 GMDIDPKKEGATVELLSPLSPDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 88
      :|||||
QY      65 SMDIDPKKEGATVELLSPLSPDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 124
      :|||||
Db      89 LCMGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMLFHISCLTFGRRETVIEYL 148
      |||
QY      125 LCMGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMLFHISCLTFGRRETVIEYL 184
      |||
Db      149 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
      |||
QY      185 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 244
      |||
Db      209 ESOC 212
      |||
QY      245 ESOC 248

RESULT  13
ENTRY   S53223
TITLE   #type complete
        e antigen precursor / core antigen - hepatitis B virus
        (isolate patient Chighine-1'85)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORANISM #formal_name hepatitis B virus, HBV
        isolate patient Chighine-1'85
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
        08-Sep-1997

ACCESSIONS S53223

```

```

REFERENCE S53112
AUTHORS Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
SUBMISSION submitted to the EMBL Data Library, March 1995
#accession S53223
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85295; NID:q736134; PID:q736136
#experimental_source isolate patient Chighine-1'85

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
        predicted #label ECP
SUMMARY #length 212 #molecular-weight 24364 #checksum 1123

Query Match 73.7%; Score 1366; DB 2; Length 212;
Best Local Similarity 97.8%; Pred. No. 1.39e-197;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db      29 GMDIDPKKEGATVELLSPLSPDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 88
      :|||||
QY      65 SMDIDPKKEGATVELLSPLSPDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 124
      :|||||
Db      89 LCMGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMLFHISCLTFGRRETVIEYL 148
      |||
QY      125 LCMGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMLFHISCLTFGRRETVIEYL 184
      |||
Db      149 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
      |||
QY      185 VSEGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 244
      |||
Db      209 ESOC 212
      |||
QY      245 ESOC 248

RESULT  14
ENTRY   S53274
TITLE   #type complete
        e antigen precursor / core antigen - hepatitis B virus
        (isolate patient Giordo'84)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORANISM #formal_name hepatitis B virus, HBV
        isolate patient Giordo'84
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
        08-Sep-1997

ACCESSIONS S53274
REFERENCE S53112
AUTHORS Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
SUBMISSION submitted to the EMBL Data Library, March 1995
#accession S53274
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85316; NID:q736208; PID:q736210
#experimental_source isolate patient Giordo'84

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
        predicted #label ECP
SUMMARY #length 212 #molecular-weight 24376 #checksum 390

Query Match 73.7%; Score 1366; DB 2; Length 212;

```

Best Local Similarity 98.9%; Pred. No. 1,39e-197;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPFSVARDLDTASALYREALSPHCSPHNTALRQAI 88
OY 65 SMDIDPYKEFGATVELLSFLPSDFPFSVARDLDTASALYREALSPHCSPHNTALRQAI 124
Db 89 LCMGELMTLATWGVNLEDPRASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 148
OY 125 LCMGELMTLATWGVNLEDPRASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 184
Db 149 VSFQWIRTPPAYRPPNAPILITLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
OY 185 VSFQWIRTPPAYRPPNAPILITLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 244
Db 209 ESQC 212
OY 245 ESQC 248

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu'89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
VARIETY isolate patient Muresu'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53169
REFERENCE S53112
#author Lal, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
##molecule_type DNA
##residues 1-183 ##label LAI
##cross-references EMBL:X85275; NID:g736057; PID:g736060
##experimental_source isolate patient Muresu'89
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular-weight 21088 #checksum 1849

Query Match 73.6%; Score 1364; DB 2; Length 183;
Best Local Similarity 98.4%; Pred. No. 3,00e-197;
Matches 180; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPFSVARDLDTASALYREALSPHCSPHNTALRQAI 60
OY 66 MDIDPYKEFGATVELLSFLPSDFPFSVARDLDTASALYREALSPHCSPHNTALRQAI 125
Db 61 CMGELMTLATWGVNLEDPRASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 120
OY 126 CMGELMTLATWGVNLEDPRASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 185
Db 121 SFGWIRTPPAYRPPNAPILITLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
OY 186 SFGWIRTPPAYRPPNAPILITLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 245
Db 181 SOC 183
OY 246 SOC 248

Search completed: Thu Dec 16 13:01:29 1999
Job time : 62 secs.

WISSETH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:01:46 1999; MasPar time 8.80 seconds
Tabular output not generated. 796.747 Million cell updates/sec

Title: >US3835-1-38183
Description: (1-248) from us3835-1-38183.pep
Perfect Score: 1854
Sequence: 1 MDIDPYKEFGATVELLSFLP.....RRRRSQSPRRRRSQSRESQC 248

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 48.750; Variance 106.745; scale 0.457

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1374	74.1	183	1	CORA_HPBV CORE ANTIGEN.	6.54e-229
2	1358	73.2	211	1	CORA_HPBV CORE ANTIGEN.	7.60e-226
3	1349	72.8	183	1	CORA_HPBV CORE ANTIGEN.	4.02e-224
4	1336	72.1	183	1	CORA_HPBV CORE ANTIGEN.	1.24e-221
5	1334	72.0	183	1	CORA_HPBV CORE ANTIGEN.	3.00e-221
6	1322	71.3	183	1	CORA_HPBV CORE ANTIGEN.	5.93e-219
7	1321	71.3	185	1	CORA_HPBV CORE ANTIGEN.	9.22e-219
8	1314	70.9	195	1	CORA_HPBV CORE ANTIGEN.	2.01e-217
9	1313	70.8	212	1	CORA_HPBV CORE ANTIGEN.	3.13e-217
10	1308	70.6	183	1	CORA_HPBV CORE ANTIGEN.	2.83e-216
11	1305	70.4	185	1	CORA_HPBV CORE ANTIGEN.	1.06e-215
12	1304	70.3	214	1	CORA_HPBV CORE ANTIGEN.	1.65e-215
13	938	50.6	188	1	CORA_HPBV CORE ANTIGEN.	5.93e-146
14	934	50.4	217	1	CORA_HPBV CORE ANTIGEN.	3.37e-145
15	933	50.3	187	1	CORA_HPBV CORE ANTIGEN.	5.20e-145
16	234	12.6	289	1	CETP_MESAT CHOLESTERYL ESTER TRAN	4.53e-19
17	234	12.6	493	1	CETP_MACRA CHOLESTERYL ESTER TRAN	4.53e-19
18	234	12.6	493	1	CETP_HUMAN CHOLESTERYL ESTER TRAN	4.53e-19
19	200	10.8	305	1	CORA_HPBH CORE ANTIGEN.	9.28e-14
20	200	10.8	305	1	CORA_HPBD CORE ANTIGEN.	9.28e-14
21	200	10.8	305	1	CORA_HPBD CORE ANTIGEN.	9.28e-14
22	200	10.8	305	1	CORA_HPBD CORE ANTIGEN.	9.28e-14
23	200	10.8	305	1	CORA_HPBD CORE ANTIGEN.	9.28e-14

ID	Accession	Standard	Prot	Length	Score	Description
24	200	10.8	497	1	CETP_RABIT CHOLESTERYL ESTER TRAN	9.28e-14
25	158	8.5	196	1	SFR2_CAEL PUTATIVE SPLICING FACT	1.43e-07
26	148	8.0	208	1	YSX2_CAEL HYPOTHETICAL 24.0 KD P	3.53e-06
27	136	7.3	78	1	PR1_SEPOF SPERMATID-SPECIFIC PRO	1.46e-04
28	131	7.1	61	1	HSP1_MACRU SPERM PROTEIN P1	6.57e-04
29	130	7.0	77	1	PR2_SEPOF SPERMATID-SPECIFIC PRO	8.85e-04
30	126	6.8	498	1	VE2_HPV08 REGULATORY PROTEIN E2	2.88e-03
31	126	6.8	1523	1	SON_HUMAN SON PROTEIN (SON3)	2.88e-03
32	124	6.7	57	1	HSP1_DIDNA SPERM PROTEIN P1	5.15e-03
33	125	6.7	60	1	HSP1_MACGI SPERM PROTEIN P1	3.85e-03
34	125	6.7	61	1	HSP1_MACRG SPERM PROTEIN P1	3.85e-03
35	122	6.6	91	1	PH1_MATED SPERM-SPECIFIC PROTEIN	9.16e-03
36	122	6.6	132	1	PR1_ANTGR SPERM HISTONE	9.16e-03
37	121	6.5	61	1	HSP1_CHICK SPERM HISTONE (PROTAMI	1.22e-02
38	121	6.5	739	1	DD13_CAEL PUTATIVE PRE-MRNA SPLI	1.22e-02
39	121	6.5	843	1	CY11_BRUMA PEPTIDYLPROLYL ISOMERA	1.22e-02
40	119	6.4	50	1	HSP1_RAT SPERM PROTEIN P1 (CY	2.15e-02
41	119	6.4	50	1	HSP1_MOUSE SPERM PROTEIN P1 (CY	2.15e-02
42	118	6.4	238	1	SFR7_HUMAN SPLICING FACTOR, ARGIN	2.86e-02
43	118	6.4	576	1	CAT1_RHOCA PEROXIDASE / CATALASE	2.86e-02
44	115	6.2	102	1	HSP2_MACRU SPERM HISTONE P2 PRECU	6.60e-02
45	115	6.2	104	1	HSP2_CALYA SPERM HISTONE P2 PRECU	6.60e-02

ALIGNMENTS

RESULT 1
ID CORA_HPBV STANDARD: PR1: 183 AA.
AC P03146; 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE AYV).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 81012091.
RA GALIBERT F., MANDART E., FITOUSSI F., TIOLLAIS P., CHARNAY P.;
RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayv)
RT cloned in E. coli.";
RL NATURE 281:646-650(1979).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PHB320).
RX MEDLINE: 85204397.
RA BICHRO V., PUSHRO P., DRELLINA D., PUMPHEN P., GREN E.;
RT "Subtype ayw variant of hepatitis B virus. DNA primary structure
RT analysis.";
RL FEBS LETT. 185:208-212(1985).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
DR EMBL: V01460; G62278; ALT.INT.
DR EMBL: X02496; -; NOT_ANNOTATED_CDS.
DR PIR: A03711; NKVLAH.
DR PIR: A03712; NKVLBH.
DR PIR: PFO0906; Hepatitis_core.1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
FT REPEAT 177 177
FT CONFLICT 33 33 T -> N (IN REF. 2).
FT CONFLICT 80 80 A -> I (IN REF. 2).
SQ SEQUENCE 183 AA; 2116 MW; 2AE7417A CRC32;
Query Match 74.1%; Score 1374; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 6.54e-229;

Matches 183: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

```
Db 1 MDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 60
   66 MDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 125
Qy 61 CMGELMTLATWGVNEDPASRDLYSVYNTNMGKLFROLWFHISCLTFGRETVIEYLV 120
   126 CMGELMTLATWGVNEDPASRDLYSVYNTNMGKLFROLWFHISCLTFGRETVIEYLV 185
Db 121 SFGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQ 180
   186 SFGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQ 245
Qy 181 SOC 183
   246 SOC 248
Db 246 SOC 248
```

RESULT 2
ID CORA_HPBVA STANDARD: PRT: 211 AA.
AC P24023:
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN ALPHA).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN PIR: A34773; NKVLAI2.
RX MEDLINE: 90266476.
RA TONG S., LI J., VIRVITSKI L., TREPO C.;
RT "Active hepatitis B virus replication in the presence of anti-HBe is
RT associated with viral variants containing an inactive pre-C region.",
RT VIROLOGY 176:556-603(1990).
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M32138; G495034; ALT_SEQ.
DR PIR: A34773; NKVLAI.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 183 190
FT REPEAT 198 206
SQ SEQUENCE 211 AA: 24208 MW: 0924CESE CRC32:

Query Match 73.2%: Score 1358; DB 1; Length 211;
Best Local Similarity 95.2%; Pred. No. 7.60e-226;
Matches 179: Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Db 24 LGMDMDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 83
   61 LGMDMDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 120
Qy 84 ROALICWGLDITLSTWGVNEDPSTSDLYSVYNTNMGKLFROLWFHISCLTFGRETV 143
   121 ROALICWGLDITLSTWGVNEDPSTSDLYSVYNTNMGKLFROLWFHISCLTFGRETV 180
Db 144 IEYVSVGWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRR 203
   181 IEYVSVGWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRR 240
Qy 204 TOSRESOC 211
   241 SOSRESOC 248
```

3.

RESULT 3
ID CORA_HPBVA STANDARD: PRT: 183 AA.
AC P03147:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADYW).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN PIR: A08967; G411874; -.
RX MEDLINE: 81012115.
RA PASEK M., GOTO T., GILBERT W., ZINK B., SCHALLER H., MCKAY P.,
RA LEADBETTER G., MURRAY K.;
RT "Hepatitis B virus genes and their expression in E. coli.",
RT NATURE 282:575-579(1979).
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: J02202; G329638; -.
DR EMBL: A08967; G411874; -.
DR PIR: B92217; NKVLAI2.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 21042 MW: 6EA6B674 CRC32:

Query Match 72.8%: Score 1349; DB 1; Length 183;
Best Local Similarity 96.7%; Pred. No. 4.02e-224;
Matches 177: Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Db 1 MDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 60
   66 MDIDPKKEGATVELLSFLPSDFPSVRLDPTASALYREALSEPHCSPHHTAL 125
Qy 61 CMGELMTLATWGVNEDPASRDLYSVYNTNMGKLFROLWFHISCLTFGRETVIEYLV 120
   126 CMGELMTLATWGVNEDPASRDLYSVYNTNMGKLFROLWFHISCLTFGRETVIEYLV 185
Db 121 SFGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQ 180
   186 SFGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQ 245
Qy 181 SOC 183
   246 SOC 248
Db 246 SOC 248
```

RESULT 4
ID CORA_HPBVA STANDARD: PRT: 183 AA.
AC P17392:
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN OKINAWA/PODM282).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN PIR: A08967; G411874; -.
RX MEDLINE: 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
RA MIYAKAWA Y., MAYUMI M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.",


```

CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isdb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; D00330.; NOT_ANNOTATED_CDS.
DR PIR; B28925; NKVLJ2.
DR PFAM; PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SO SEQUENCE 183 AA; 21081 MW; 115BD9E3 CRC32;

Query Match          72.1%; Score 1336; DB 1; Length 183;
Best Local Similarity 96.2%; Pred. No. 1.24e+22;
Matches 176; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db      1 MDIDYKEFGASVSELLSLPSDFSPSVRDLDLTASALYREALSEPHKSPHHTARQAIL 60
        |||
Oy      66 MDIDYKEFGATVELLSLPDFFPVSVDRLDTASALYREALSEPHKSPTHTRQLRQAIL 125
        |||
Db      61 CMGELMNALTWVGSLLEDPASRELIVSYVNVMGAKIQOLMFHSICLTGFGEITYLEYLY 120
        |||
Oy      126 CMGELMTLATWGVGLLEDPASRDLVASYNTMGMKIFROLMLFHSICLTGFGETIEXLY 185
        |||
Db      121 SFGVVIARTPPAPRPNPAPILSTLPETTVRRGRSPPRRTPSPRRRSOSPRRRSSORE 180
        |||
Oy      186 SFGVIRIPRAPRPNPAPILSTLPETTYVRGRSRPRRTPTSPRRRSOSPFRRSSORE 245
        |||
Db      181 SOC 183
        |||
Oy      246 SOC 248

RESULT    5 STANDARD:   FRT: 183 AA.
ID CORA_HPBV4
AC P03150; P03151;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADP), HEPATITIS B VIRUS (SUBTYPE ADRA4),
OS AND HEPATITIS B VIRUS (SUBTYPE ADM / STRAIN INDONESIA/PIDMA20).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ADR;
RX MEDLINE: 83168919.
RA ONO Y., OSAIDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.:
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA: subtype adr and adw."
RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ADR4;
RX MEDLINE: 83246570.
RA FUJIYAMA A., MIYANOHIRA A., NOZANKI C., YONEYAMA T., OTOMO N.,
RA MATSUUBARA K.:
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT adr."
RL NUCLEIC ACIDS RES. 11:4601-4610(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ADM;
RX MEDLINE: 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSOWEMINGNO R.I., IMAI M.,
RA MIYAKAWA Y., MAYUMI M.;
```

```

RT      "Typing hepatitis B virus by homology in nucleotide sequence:
rt      Comparison of surface antigen subtypes.",
RL      J. GEN. VIROL. 69:2575-2583(1988).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
cc      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; V00867; -. NOT_ANNOTATED_CDS.
DR      EMBL; X01587; G59407; -.
DR      EMBL; D000331; -. NOT_ANNOTATED_CDS.
DR      PIR; A93480; NKVLAS.
DR      PIR; B93460; NKVLAA.
DR      PIR; C28925; NKVLJ3.
DR      PIR; PF00906; Hepatitis_core.1.
KW      CORE PROTEIN; REPEAT.
FT      REPEAT      162      169
FT      REPEAT      170      177
SQ      SEQUENCE      183 AA; 21095 MW; D2E185B9 CRC32;
Query Match      72.0%; Score 1334; DB 1; Length 183;
Best Local Similarity 95.6%; Pred. No. 3,00e-221;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Db      1 MDIDPYKEFGASVELLSFLPSDFPSIRDLDTASALYREALLESPHCHSPHTALRQAIL 60
Qy      66 MDIDPYKEFGATVELLSFLPSDFPSVADLDDTASALYREALLESPHCHSPHTALRQAIL 125
Db      61 CWGELMNLATWGSLEDPASRELVSIVNNMGKIQQLMFHISCTLFGRETYLEIY 120
Qy      126 CWGELMTATATWGVNLLEDPASRDLYSVVNNMGKIFQLMFHISCTLFGRETYIEIY 185
Db      121 SFGWIRTTPPYRPPRNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSSRE 180
Qy      186 SFGWIRTTPPYRPPRNAILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSSRE 245
Db      181 SQC 183
Qy      246 SQC 248
RESULT      6
ID      CORA_HPBVT      STANDARD;      PRT;      183 AA.
AC      P17391.
DT      01-AUG-1980 (REL. 15, CREATED)
DT      01-AUG-1980 (REL. 15, LAST SEQUENCE UPDATE)
DT      01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE      CORE ANTIGEN.
OS      HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN JAPAN/PJDM233).
OC      VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN      (1)
RP      SEQUENCE FROM N.A.
RX      MEDLINE; 89010694.
RA      OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
RA      MIYAKAWA Y., MAYUMI M.;
RT      "Typing hepatitis B virus by homology in nucleotide sequence:
RT      Comparison of surface antigen subtypes.";
RL      J. GEN. VIROL. 69:2575-2583(1988).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D00329; -. NOT_ANNOTATED_CDS.
DR      PIR; A28925; NKVLJ1.

```

DR PFAM: PF00906; Hepatitis_core; 1.
 KW CORE PROTEIN; REPEAT.
 FT REPEAT 162 169
 FT REPEAT 170 177
 SQ SEQUENCE 183 AA; 21224 MW; F6B348B6 CRC32;

Query Match
 Best Local Similarity 96.2%; Pred. No. 5,93e-219;
 Matches 176; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPYKEGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAIL 60
 |||||||
 QY 66 MDIDPYKEGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAIL 125
 |||||||
 Db 61 CWGELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 120
 |||||||
 QY 126 CWGELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 185
 |||||||
 Db 121 SFGWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 180
 |||||||
 QY 186 SFGWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 245
 |||||||
 Db 181 SOC 183
 |||
 QY 246 SOC 248

RESULT 7
 ID CORA_HPBVT STANDARD; PRT; 185 AA.
 AC P03178;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C
 OS HEPATITIS B VIRUS (SUBTYPE ADW).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 83168919.
 RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGIMO Y., NISHIOKA K.;
 RT "The complete nucleotide sequences of the cloned hepatitis B virus
 RT DNA; subtype adw and adw.";
 RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: V00866; NOT_ANNOTATED_CDS.
 DR PIR: C93460; NKVLAC.
 DR PFAM: PF00906; Hepatitis_core; 1.
 KM CORE PROTEIN; REPEAT.
 FT REPEAT 164 171
 FT REPEAT 172 179
 SQ SEQUENCE 185 AA; 21394 MW; 791E0381 CRC32;

Query Match
 Best Local Similarity 96.2%; Pred. No. 9.22e-219;
 Matches 178; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Db 1 MDIDPYKEGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAIL 60
 |||||||
 QY 66 MDIDPYKEGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAIL 125
 |||||||
 Db 61 CWGELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 120
 |||||||
 QY 126 CWGELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 185
 |||||||

Db 121 SFGWIRTPPAYRPNPAPILSTLPETTVVRRDRGSPRRRTSPRRRSQSPRRRSQSPRE 180
 |||||||
 QY 186 SFGWIRTPPAYRPNPAPILSTLPETTVRR--GRSPRRRTSPRRRSQSPRRRSQSPRE 243
 |||||||
 Db 181 RESOC 185
 |||||
 QY 244 RESOC 248

RESULT 8
 ID CORA_HPBVT STANDARD; PRT; 195 AA.
 AC P29178;
 DT 01-DEC-1992 (REL. 24, CREATED)
 DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C
 OS HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90169850.
 RA BHAT R.A., ULRICH P.P., VYAS G.N.;
 RT "Molecular characterization of a new variant of hepatitis B virus in
 RT a persistently infected homosexual man.";
 RL HEPATOLOGY 11:271-276(1990).
 DR PIR: A37182; NKVLH3.
 DR PFAM: PF00906; Hepatitis_core; 1.
 KM CORE PROTEIN; REPEAT.
 FT REPEAT 174 181
 FT REPEAT 182 189
 SQ SEQUENCE 195 AA; 22461 MW; AF3DB5F3 CRC32;

Query Match
 Best Local Similarity 93.5%; Pred. No. 2.01e-217;
 Matches 173; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 11 FGDLIDPYKEGATVLLSFLPSDFPSVRDLDTASALYRESLESSHCSPHHTALROA 70
 |||||||
 QY 64 FGDLIDPYKEGATVLLSFLPSDFPSVRDLDTASALYRESLESSHCSPHHTALROA 123
 |||||||
 Db 71 ILGWELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 130
 |||||||
 QY 124 ILGWELMTLATVGNLDPASRDLYVNVNTNMGKIRQLMFIISCLTGRETVIEYLV 183
 |||||||
 Db 131 LVSGWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 190
 |||||||
 QY 184 LVSGWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 243
 |||||||
 Db 191 RESOC 195
 |||||
 QY 244 RESOC 248

RESULT 9
 ID CORA_HPBVT STANDARD; PRT; 212 AA.
 AC Q05495;
 DT 01-NOV-1995 (REL. 32, CREATED)
 DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C
 OS HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE W4B).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93346970.
 RA NAUMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,
 RA GERLICH W.H.;
 RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
 RT that expresses HBV surface antigen subtype adw4.";
 RL J. GEN. VIROL. 74:1627-1632(1993).
 CC

```

CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC
CC      EMBL; X69798; G59423; -.
DR      PFAM; PF00906; Hepatitis_core; 1.
KW      CORE PROTEIN; REPEAT.
FT      DOMAIN 178 204 ARG-RICH.
FT      REPEAT 184 203 3 x 5 AA REPEATS OF S-P-R-R-R.
FT      REPEAT 184 188 1.
FT      REPEAT 191 195 2.
FT      REPEAT 199 203 3.
SQ      SEQUENCE 212 AA; 24234 MW; BCCDF263 CRC32;
Query Match 70.8%; Score 1313; DB 1; Length 212;
Best Local Similarity 92.9%; Pred. No. 3,13e-217;
Matches 17; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
Db 29 GMDIDPYKEFGASVLLFLPDEFPYSVRLDLDTSALYRLDLESPECHPTHTALROAI 88
Oy 65 SMDIDPYKEFGATVLLSTFLSPDFPSVRLDLDTSALYRLDLESPECHSPHTALROAI 124
Db 89 LCMGLMTLASVWGNLDEDPARDLVVWVYVNTNMGILKIRQLMFIISCLTGFRETVLYL 148
Oy 125 LCMGLMTLATVGVGNLDEDPASRLDLYSVYVNTNMGILKIRQLMFIISCLTGFRETVLYL 184
Db 149 VSFQGWITPPAYRPPNAPILSTLDETTVVRNRGSPRRRTSPRRRSQSPRRRSOSP 208
Oy 185 VSFQGWITPPAYRPPNAPILSTLDETTVVRNRGSPRRRTSPRRRSQSPRRRSOSP 244
Db 209 ASOC 212
Oy 245 ESOC 248
RESULT 10 STANDARD: PRT: 183 AA.
AC CORA.HPBVL
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN LSH / CHIMPANZEE ISOLATE).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
OC (1)
RX SEQUENCE FROM N.A.
RX MEDLINE: 88258473.
RA VAUDIN M., WOLSTENHOLME A.J., TSIOUANE K.N., ZUCKERMAN A.J.,
RA HARRISON F.J.;
RT "The complete nucleotide sequence of the genome of a hepatitis B
RT virus isolated from a naturally infected chimpanzee.";
RL J. GEN. VIROL. 69:1383-1389(1988).
CC
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC
CC      EMBL; D00220; D1000603; -.
DR      PIR; A28885; NKVLCF.
DR      PFAM; PF00906; Hepatitis_core; 1.
KW      CORE PROTEIN; REPEAT.
FT      REPEAT 162 169
FT      REPEAT 170 177
SQ      SEQUENCE 183 AA; 20999 MW; IAF57C9 CRC32;

```

Query Match	70.6%	Score 1308	DB 1	Length 183
Best Local Similarity	94.5%	Pred. No. 2.8e-216		
Matches 173	Conservative 6	Mismatches 4	Indels 0	Gaps 0
Db	1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSPHCSPHNTALROAIL 60			
Qy	66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSPHCSPHNTALROAIL 125			
Db	61 CWMGLMTLATVAVGNLLEDPASRDLVSVYNTNMGKLRQLMLFHIISCLTGRETVLEYLV 120			
Qy	126 CWMGLMTLATVAVGNLLEDPASRDLVSVYNTNMGKLRQLMLFHIISCLTGRETVLEYLV 185			
Db	121 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPA 180			
Qy	186 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 245			
Db	181 SQC 183			
Qy	246 SQC 248			
RESULT 11				
ID	CORA_HPBV2	STANDARD	PRT	185 AA.
AC	P03148			
DT	21-JUL-1986 (REL. 01, CREATED)			
DT	01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			
OC	HEPATITIS B VIRUS (SUBTYPE ADM2).			
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	VALENUEUA P., QUIROGA M., ZALDIVAR J., GRAY P., RUTTER W.J.,			
RL	(IN) FIELD B.N., JAENISCH R., FOX C.F. (EDS.);			
RL	ANIMAL VIRUS GENETICS, PP.57-70, ACADEMIC PRESS, NEW YORK (1980).			
DR	PIR: A94409: NKYL3.			
DR	PFAM: PF00906; Hepatitis-core: 1.			
KM	CORE PROTEIN; REPEAT.			
FT	REPEAT 164 171			
FT	REPEAT 172 179			
SO	SEQUENCE 185 AA; 21304 MW; E2EA3360 CRC32;			
Query Match	70.4%	Score 1305	DB 1	Length 185
Best Local Similarity	95.7%	Pred. No. 1.0e-215		
Matches 177	Conservative 4	Mismatches 2	Indels 2	Gaps 1
Db	1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSPHCSPHNTALROAIL 60			
Qy	66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSPHCSPHNTALROAIL 125			
Db	61 CWMGLMTLATVAVGNLLEDPASRDLVSVYNTNMGKLRQLMLFHIISCLTGRETVLEYLV 120			
Qy	126 CWMGLMTLATVAVGNLLEDPASRDLVSVYNTNMGKLRQLMLFHIISCLTGRETVLEYLV 185			
Db	121 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPA 180			
Qy	186 SFGWIRPPAIRPPAPNPAILSTLLETIVVRR--GRSPRRRTSPRRRSQSPRRRSQSPA 243			
Db	181 RESOC 185			
Qy	244 RESOC 248			
RESULT 12				
ID	CORA_HPBV9	STANDARD	PRT	214 AA.
AC	P17099			
DT	01-AUG-1990 (REL. 15, CREATED)			
DT	01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)			
DT	01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			

Query Match	70.6%	Score 1308	DB 1	Length 183
Best Local Similarity	94.5%	Pred. No. 2.8e-216		
Matches 173	Conservative 6	Mismatches 4	Indels 0	Gaps 0
Db	1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSEPHCSPHNTALROAIL 60			
Qy	66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSEPHCSPHNTALROAIL 125			
Db	61 CWMGLMTLATVAVGNLLEDPASRDLYVSVNTNMGKLRQLMLFHIISCLTGRETVLEYLV 120			
Qy	126 CWMGLMTLATVAVGNLLEDPASRDLYVSVNTNMGKLRQLMLFHIISCLTGRETVLEYLV 185			
Db	121 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPA 180			
Qy	186 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 245			
Db	181 SQC 183			
Qy	246 SQC 248			
RESULT	11			
ID	CORA_HPBV2	STANDARD	PRT	185 AA.
AC	P03148			
DT	21-JUL-1986 (REL. 01, CREATED)			
DT	01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			
OC	HEPATITIS B VIRUS (SUBTYPE ADM2).			
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	VALENUEUA P., QUIROGA M., ZALDIVAR J., GRAY P., RUTTER W.J.,			
RL	(IN) FIELD B.N., JAEINISCH R., FOX C.F. (EDS.);			
RL	ANIMAL VIRUS GENETICS, PP.57-70, ACADEMIC PRESS, NEW YORK (1980).			
DR	PIR: A94409: NKYL3.			
DR	PFAM: PF00906; Hepatitis-core: 1.			
KM	CORE PROTEIN; REPEAT.			
FT	REPEAT 164 171			
FT	REPEAT 172 179			
SO	SEQUENCE 185 AA; 21304 MW; E2EA3360 CRC32;			
Query Match	70.4%	Score 1305	DB 1	Length 185
Best Local Similarity	95.7%	Pred. No. 1.0e-215		
Matches 177	Conservative 4	Mismatches 2	Indels 2	Gaps 1
Db	1 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSEPHCSPHNTALROAIL 60			
Qy	66 MDIDPYKEFGATVELLSFLPSDFPSVRDLDDTASALYREALSEPHCSPHNTALROAIL 125			
Db	61 CWMGLMTLATVAVGNLLEDPASRDLYVSVNTNMGKLRQLMLFHIISCLTGRETVLEYLV 120			
Qy	126 CWMGLMTLATVAVGNLLEDPASRDLYVSVNTNMGKLRQLMLFHIISCLTGRETVLEYLV 185			
Db	121 SFGWIRPPAIRPPAPNPAILSTLLETIVVRRGRSPRRRTSPRRRSQSPRRRSQSPA 180			
Qy	186 SFGWIRPPAIRPPAPNPAILSTLLETIVVRR--GRSPRRRTSPRRRSQSPRRRSQSPA 243			
Db	181 RESOC 185			
Qy	244 RESOC 248			
RESULT	12			
ID	CORA_HPBV9	STANDARD	PRT	214 AA.
AC	P17099			
DT	01-AUG-1990 (REL. 15, CREATED)			
DT	01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)			
DT	01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			


```
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on lists  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isib.ch/announce/  
CC or send an email to license@isib-stb.ch).  
CC -----  
DR EMBL; K02715; G325401; -.  
DR PIR; A03715; NKVLS.  
KW CORE PROTEIN; REPEAT. 1.  
FT REPEAT 196 203  
FT REPEAT 204 211  
SQ SEQUENCE 217 AA: 25189 MW: B70A00B1 CRC32;  
  
Query Match 50.4%; Score 934; DB 1; Length 217;  
Best Local Similarity 68.1%; Pred. No. 3,37e-145;  
Matches 128; Conservative 22; Mismatches 32; Indels 6; Gaps 2;  
  
Db 31 MDIDYKEFGSSGYOLNPLPDLFPFDNALVDTAALVEEELTGDEHCSPHHTAIRQALV 90  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Oy 66 MDIDYKEFGGAVELSLFSPDFPSVRDLDLTASALYLEALESEBHCHSHTTALAQQAIL 125  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 91 CMEELTRITWMSKNTETEVRR-IIVDHNHWGSKVGRTLMFHLSCITFGOTVOEFLY 149  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Oy 126 CGEMLMTLATWGVLVEDASRDVLVSIVNYTMGLFRQLMFLHSCLTFGEHYIEYLTV 185  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 150 SGFWIIRTAPYPAPPNPAPILSTLPPEHTVIARRGGSRRAASPRRTPSPRRRSQSPPRRR 209  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Oy 186 SGFWIIRTAPPAARPAPNPAPILSTLPPEHTVRRRG-----RSPRRTRTPSPRRRSQSPPRRR 240  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Db 210 SOSPASNC 217  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
Oy 241 SOSRESOC 248  
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :  
  
RESULT 15  
ID CORA_MHV8 STANDARD: PRT: 187 AA.  
AC P06433:  
DT 01-JAN-1988 (REL. 06, CREATED)  
DD 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)  
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)  
DE CORE ANTIGEN.  
GN C.  
GC WOODCHUCK HEPATITIS VIRUS 8 (WHV 8).  
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE: 86062931.  
RA KODAMA K., OGASAWARA N., YOSHIKAWA H., MURAKAMI S.;  
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:  
RL J. VIROL. 56:978-986(1985)."  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-stb.ch/announce/  
CC or send an email to license@isb-stb.ch).  
CC -----  
DR EMBL; M11082; G336135; -.  
DR PIR; A03714; NKVIC2.  
KW CORE PROTEIN; REPEAT.  
FT REPEAT 166 173  
FT REPEAT 174 181  
SQ SEQUENCE 187 AA: 21579 MW: A1C354F3 CRC32;  
  
Query Match 50.3%; Score 933; DB 1; Length 187;  
Best Local Similarity 66.3%; Pred. No. 5,20e-145;  
Matches 124; Conservative 26; Mismatches 33; Indels 4; Gaps 2;
```

```

Db      1 MDIDPYKEFGSSYOLLNLELPDFFPPDNLAVDTATLYEEELTGREGCSPHHAIROALV 60
Qy      66 MDIDPYKEFGATVELLSLTFPSDFPSPYRDLDITASALYKRELBSPBECSSHHHAIKRAIL 125
Db      61 CWDELTKLIAMSSNITSSEOVRTIIVNVHNDTWGLKVKROSIMFHLSCLTFGQHTVOEFLV 120
Qy      126 CWGELMTLATVWGVNLEDDPASRDLVSVMVTNMGLKFRQLLMFHSICLTFGRETVIEYL 165
Db      121 SFVVMIRPAPYRRPPNAPILSTLEHVIYIRGGKARASRPKRRTPSPRRRRSSQSPRRRRS 180
Qy      186 SFGVWIRPAPYRRPPNAPILSTLEHVIYVR--RG-RSPRRRTSPSPRRRRSSQSPRRRRS 241
Db      181 QSPSANC 187
Qy      242 QSRESQC 248

```

Search completed: Thu Dec 16 13:02:25 1999
Job time : 39 secs.

 WISE (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Thu Dec 16 13:02:42 1999; MasPar time 18.20 Seconds
 Tabular output not generated. 743.783 Million cell updates/sec

Title: >US3835-1-38183
 Description: (1-248) from US3835-1-38183.pep
 Perfect Score: 1854
 Sequence: 1 MDIDPKKFGATVLLSLP.....RRRSOSPRRRRSQSRSSQC 248

Scoring table: PAM 150
 Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: spltemb19
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_orfanelle
 9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 46.830; Variance 106.555; scale 0.439

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	74.2	212 14	089656	PRE-C/CORE.	2.78e-221
2	1373	74.1	183 14	089437	X, PREC AND C GENES (C	6.51e-221
3	1373	74.1	183 14	068008	X, PREC AND C GENES (F	6.51e-221
4	1374	74.1	212 14	068020	PRE-C/CORE.	4.26e-221
5	1374	74.1	212 14	089597	HBCAG.	4.26e-221
6	1373	74.1	212 14	067876	PRE-C/C ORF.	6.51e-221
7	1372	74.0	212 14	011884	CORE ANTIGEN PRECURSOR	9.96e-221
8	1370	73.9	183 14	068066	X, PREC AND C GENES (L	2.33e-220
9	1370	73.9	212 14	068025	PRE-C/CORE.	2.33e-220
10	1369	73.9	212 14	068025	PRE-C/C ORF.	3.57e-220
11	1369	73.8	212 14	067984	PRE-C/CORE.	3.57e-220
12	1369	73.8	212 14	068032	PRE-C/CORE.	3.57e-220
13	1368	73.8	212 14	068010	PRE-C/CORE.	5.46e-220
14	1368	73.7	212 14	068070	PRE-C/CORE.	1.28e-219
15	1366	73.7	212 14	068030	PRE-C/CORE.	1.28e-219
16	1364	73.6	183 14	067989	X, PREC AND C GENES (M	2.99e-219
17	1364	73.5	183 14	068048	X, PREC AND C GENES (F	4.57e-219
18	1363	73.5	212 14	068014	PRE-C/CORE.	4.57e-219
19	1363	73.5	212 14	068012	PRE-C/CORE.	4.57e-219
20	1363	73.5	212 14	068012	PRE-C/CORE.	4.57e-219

Result ID	1	PRELIMINARY:	PRT:	212 AA.
AC	089656			
AD	089656			
DT	01-NOV-1996 (TREMREL. 01, CREATED)			
DT	01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)			
DE	PRE-C/CORE.			
GN	PRE-C/CORE.			
OS	HEPATITIS B VIRUS.			
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RN	1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-PATIENT LICHEN-1'85;			
RA	LAI M.E., MAZOLENTI A.P., PORRU A., BALESTRIERI A.;			
RL	SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	12]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATM4;			
RA	PLUCIENNICZAK A.;			
RL	SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	13]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 81012091.			
RA	GALIBERT F., MANDART E., FITOUSSI F., TIOUAI P., CHARNAY P.;			
RT	"Nucleotide sequence of the hepatitis B virus genome (subtype ayw)			
RT	cloned in E. coli.;"			
RL	NATURE 281:646-650(1979).			
RN	14]			
RP	SEQUENCE FROM N.A.			
RA	BORSOVA G.P., PUMPER P.P., BYCHKO V.V., PUSHKO P.M., KALIS Y.V.,			
RA	DISHLER A.V., GREN E.Y., TSIBINOGIN V.V., KUKAIN R.A.;			
RL	DOU. BIOCHEM. 279:386-390(1985).			
DR	EMBL: X85290; G736116; -			
DR	EMBL: X85300; G736152; -			
DR	EMBL: X85313; G736196; -			
DR	EMBL: X85283; G736090; -			
DR	EMBL: X85306; G736174; -			
DR	EMBL: X85316; G527437; -			
DR	EMBL: J02203; G329642; -			
DR	EMBL: X85312; G736193; -			
DR	PFAM: PF00906; Hepatitis_core; 1.			
KW	SIGNAL.			
SO	SEQUENCE	212 AA;	24350 MW;	71EAC082 CRC32;

ALIGNMENTS

Query Match 74.2%; Score 1375; DB 14; Length 212;
Best Local Similarity 99.5%; Pred. No. 2.78e-221;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 88
QY :|||||
65 SMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 124
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 148
QY :|||||
125 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 184
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 208
QY :|||||
185 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 244
Db 209 ESOC 212
QY :|||||
245 ESOC 248

RESULT 2 PRELIMINARY; PRT; 183 AA.
ID 089437
AC 089437:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X. PREC AND C GENES (CASTA 2).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC LAI M.E., MAZZOLEN A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AW;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85289; G736113; -;
DR EMBL: X80925; E198085; -;
DR PFAM: PF00906; Hepatitis-core; 1.
DR SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 74.1%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 6.51e-221;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 60
QY :|||||
66 MDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 125
Db 61 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 120
QY :|||||
126 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 185
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 180
QY :|||||
186 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 245
Db 181 SOC 183
QY :|||||
246 SOC 248

RESULT 3

ID 068008 PRELIMINARY; PRT; 183 AA.

AC 068008:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X. PREC AND C GENES (FERRACUTI 1).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-1'89;
RL LAI M.E., MAZZOLEN A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85284; G736098; -;
DR PFAM: PF00906; Hepatitis-core; 1.
DR SEQUENCE 183 AA; 21102 MW; BB9E9AFB CRC32;

Query Match 74.1%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 6.51e-221;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 60
QY :|||||
66 MDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 125
Db 61 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 120
QY :|||||
126 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 185
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 180
QY :|||||
186 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 245
Db 181 SOC 183
QY :|||||
246 SOC 248

RESULT 4 PRELIMINARY; PRT; 212 AA.

ID 068020
AC 068020:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG-1'85;
RL LAI M.E., MAZZOLEN A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85291; G736119; -;
DR PFAM: PF00906; Hepatitis-core; 1.
DR SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 74.1%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4.26e-221;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 88
QY :|||||
65 SMDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSPHCSPHHTALROAI 124
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 148
QY :|||||
125 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVEYL 184
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 208
QY :|||||
185 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSOSPRRRRSOSR 244

Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 5
ID 089597 PRELIMINARY: PRT: 212 AA.
AC 089597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HECAG.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SUB-TYPE AYW;
RX MEDLINE: 94079539.
RA PREISLER-ADAMS S., SCHLAYER M.J., PETERS T., HETTLER F., GEROK W.,
RA RASENACK J.;
RT "sequence analysis of hepatitis B virus DNA in immunologically
negative infection";
RL ARCH. VIROL. 133:385-396(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X72702; G288930; -;
DR EMBL: X80925; E198084; -;
DR PRAM: PF00906; Hepatitis_core: 1;
SQ SEQUENCE 212 AA; 24336 MW; 575020BF CRC32;

Query Match 74.1%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4,26e-221;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 88
:|||||
QY 65 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 124

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 148
:|||||
QY 125 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 184

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 185 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 244

Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 6
ID 067876 PRELIMINARY: PRT: 212 AA.
AC 067876;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CI, HBV SUBTYPE AYW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65258; G59436; -;
DR PRAM: PF00906; Hepatitis_core: 1;
SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;

Query Match 74.1%; Score 1373; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 6,51e-221;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 88
:|||||
QY 65 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 124

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 148
:|||||
QY 125 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 184

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 185 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 244

Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 7
ID 011884 PRELIMINARY: PRT: 212 AA.
AC 011884;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN PRECURSOR.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA RAO B.S., CASEY J.L., RINAUDO J.S., KOREA B.E.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U95551; G2182120; -;
DR PRAM: PF00906; Hepatitis_core: 1;
SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;

Query Match 74.0%; Score 1372; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 9,96e-221;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 88
:|||||
QY 65 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALESPHCSPHHTALRQAI 124

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 148
:|||||
QY 125 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLIMFHSICLTFGRTVIEYL 184

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 185 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 244

Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 8
ID 068066 PRELIMINARY: PRT: 183 AA.
AC 068066;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X. PREC AND C GENES (LICHERI 2).
GN CORE.
OS HEPATITIS B VIRUS.

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2.87;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 6F38A3B CRC32;

Query Match 73.9%; Score 1370; DB 14; Length 183;
Best Local Similarity 98.4%; Pred. No. 2.33e-220;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 60
|||
QY 66 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 125
Db 61 CWDGLMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 120
|||
QY 126 CWDGLMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 185
Db 121 SEGVWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 180
|||
QY 186 SEGVWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 245
Db 181 SOC 183
|||
QY 246 SOC 248

RESULT 9
ID 068068 PRELIMINARY; PRT: 212 AA.
AC 068068;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3.90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;

Query Match 73.9%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.33e-220;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 88
|||
QY 65 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 124
Db 89 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 148
|||
QY 125 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 184
Db 149 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
|||
QY 185 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 244
Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 10
ID 068025 PRELIMINARY; PRT: 212 AA.

AC 068025;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG/3;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;

Query Match 73.9%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.33e-220;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 88
|||
QY 65 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 124
Db 89 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 148
|||
QY 125 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 184
Db 149 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
|||
QY 185 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 244
Db 209 ESOC 212
|||
QY 245 ESOC 248

RESULT 11
ID 067872 PRELIMINARY; PRT: 212 AA.
AC 067872;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C. HBSAG POSITIVE, HBV SUBTYPE AYW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24348 MW; A6BD9DA1 CRC32;

Query Match 73.8%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 3.57e-220;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 88
|||
QY 65 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYRDALSPHCSPHHTALROAI 124
Db 89 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 148
|||
QY 125 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIHISCLTFGRETIVIELY 184
Db 149 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
|||
QY 185 VSFQWIRTPPAYRPAPNAILSTLPETTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 244
Db 209 ESOC 212

QY	125	LCWELMTLATWGVNLEDPASRLDVSYYNTNGLKFLRQLMFWHISCLTFGRETVEYL	184
Db	149	VSFQWVIRTPPAPYPPNAPILSTLPETTVVRRKRSRRRTPSPRRRRSSPPRRRSQSR	208
QY	185	VSFQWVIRTPPAPYPPNAPILSTLPETTVVRRKRSRRRTPSPRRRRSSPPRRRSQSR	244
Db	209	ESOC 212	
QY	245	ESOC 248	
RESULT	14		
ID	068010	PRELIMINARY;	PRT; 212 AA.
AC	068010;		
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)		
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)		
DE	PRE-C/CORE.		
DE	PRE-C/CORE.		
OS	HEPATITIS B VIRUS.		
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-PATIENT FERRACUTI-2'90;		
RA	LAI M.E., MAZOLENT A.P., PORRU A., BALESTRERI A.;		
RL	SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.		
DR	EMBL; X85285; G736100; .		
SR	PFAM; PF00906; Hepatitis_core.1;		
SO	SEQUENCE 212 AA; 24292 MW; 04A4D12D CRC32;		
Query Match		73.8%; Score 1368; DB 14; Length 212;	
Best Local Similarity	98.4%;	Pred. No. 5.46e-220;	
Matches	181; Conservative	2; Mismatches	1; Indels 0; Gaps 0;
Db	29	GMDIDPYKEFGATVELLSFLPSAFPPSVRLDLPASALYREALESPHCSPHNTALRQAI	88
QY	65	SMDDIPYKEFGATVELLSFLPSDFPFSVRLDLPASALYREALESPHCSPHNTALRQAI	124
Db	89	LCWGDMLATWGVNLEDPASRLDVSYYNTNGLKFLRQLMFWHISCLTFGRETVEYL	148
QY	125	LCWELMTLATWGVNLEDPASRLDVSYYNTNGLKFLRQLMFWHISCLTFGRETVEYL	184
Db	149	VSFQWVIRTPPAPYPPNAPILSTLPETTVVRRKRSRRRTPSPRRRRSSPPRRRSQSR	208
QY	185	VSFQWVIRTPPAPYPPNAPILSTLPETTVVRRKRSRRRTPSPRRRRSSPPRRRSQSR	244
Db	209	ESOC 212	
QY	245	ESOC 248	
RESULT	15		
ID	068070	PRELIMINARY;	PRT; 212 AA.
AC	068070;		
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)		
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)		
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)		
DE	PRE-C/CORE.		
DE	PRE-C/CORE.		
OS	HEPATITIS B VIRUS.		
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-PATIENT GIORDO'84;		
RA	LAI M.E., MAZOLENT A.P., PORRU A., BALESTRERI A.;		
RL	SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.		
DR	EMBL; X85316; G736210; .		
SR	PFAM; PF00906; Hepatitis_core.1;		
SO	SEQUENCE 212 AA; 24376 MW; 80F52D0F CRC32;		
Query Match		73.7%; Score 1366; DB 14; Length 212;	
Best Local Similarity	98.4%;	Pred. No. 1.28e-219;	
Matches	182; Conservative	1; Mismatches	1; Indels 0; Gaps 0;

US3835-10-38183
MDIDPYKEFGATVELLSFLPSDFPSPVRLDLDITASLLQMDFGPEHLLVDFLOSLSMDIDPYKEFGATV
ELLSFLPSDFPSPVRLDLDITASALYREALSEPHCSPHHTALROAILCWMGLMTLATWGVNLEDDPASRD
LVSVYNTNMGLKFRQLMFMHISCLTFGRETVEYLVSGVWIRTPPAYRPPNAPILSTLPETTIVRRRG
RSPRRRTSPRRRRSQSPRRRRSQSRESQCI

 M P I S E I F
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:08:54 1999; Maspar time 11.54 Seconds

Tabular output not generated. 442.105 Million cell updates/sec

Title: >US3835-10-38183
 Description: (1-240) from us3835-10-38183.pep
 Perfect Score: 1796
 Sequence: 1 MDIDPYKFGATVELLSFLP.....RRRSQSPRRRRSQSRESQC 240

Scoring table:
 PAM 150
 Gap 11

Searched: 170751 segs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.355; Variance 174.564; scale 0.191

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	76.6	194	32	W50242	Hepatitis B virus pre
2	1375	76.6	212	32	W50250	Hepatitis B virus pre
3	1375	76.6	346	5	R27473	S12/core protein.
4	1363	75.9	184	1	P80959	Hepatitis B virus sub
5	1359	75.7	193	32	W50241	Hepatitis B virus pre
6	1358	75.6	183	32	W50251	Hepatitis B virus p21
7	1353	75.3	183	20	W09044	Hepatitis B virus cor
8	1349	75.1	183	1	R05635	Hepatitis B antigen.
9	1349	75.1	183	5	P00041	Sequence of core anti
10	1344	74.8	397	20	W09048	Plasmodium falciparum
11	1338	74.5	184	5	P00004	Sequence of core anti
12	1336	74.4	196	8	R40806	Hepatitis B core / PV
13	1337	74.4	208	8	R40808	Hepatitis B core / PV
14	1334	74.3	183	3	P40311	Hepatitis B virus core
15	1334	74.3	183	8	R40805	Hepatitis B core prot
16	1334	74.3	183	13	R68868	Hepatitis B virus pol

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description	Pred. No.
1	W50242	74.3	183	12	R62869	Hepatitis B virus cor	1.29e-95
2	W50242	74.2	198	8	R40807	Hepatitis B core / IL	1.55e-95
3	W50242	73.9	293	10	R55286	Prest full length cor	3.96e-95
4	W50242	73.8	183	13	R68866	Hepatitis B virus pol	6.94e-95
5	W50242	73.8	214	1	P90702	Deduced amino acid se	6.94e-95
6	W50242	73.8	289	20	W09049	Plasmodium falciparum	6.94e-95
7	W50242	73.7	185	6	R30861	Hepatitis B core anti	8.36e-95
8	W50242	73.7	185	10	R55284	Deduced sequence of f	8.36e-95
9	W50242	73.7	185	6	R30780	Native HBcAg protein.	8.36e-95
10	W50242	73.6	185	26	P30061	Core antigen of HBV.	1.47e-94
11	W50242	73.5	183	13	R68869	Hepatitis B virus pol	1.77e-94
12	W50242	73.3	185	6	R30781	Altered HBcAg protein	3.10e-94
13	W50242	73.3	185	6	R31026	Altered HBcAg protein	3.10e-94
14	W50242	73.2	183	13	R68867	HBcAg/Beta-Gal fusion	5.43e-94
15	W50242	73.1	183	13	R68870	Hepatitis B virus pol	5.54e-94
16	W50242	73.0	185	6	R30784	Altered HBcAg protein	9.51e-94
17	W50242	73.0	185	6	R31029	Altered HBcAg protein	9.51e-94
18	W50242	73.0	185	6	R30865	Modified Hepatitis B	9.51e-94
19	W50242	72.9	185	6	R30862	Hepatitis B core anti	1.38e-93
20	W50242	72.8	185	6	R31028	Altered HBcAg protein	1.67e-93
21	W50242	72.8	185	6	R30864	Modified Hepatitis B	1.67e-93
22	W50242	72.6	214	1	P80961	HBV core antigen enco	4.24e-93
23	W50242	72.5	203	1	P82872	HBcAg/Beta-Gal fusion	5.12e-93
24	W50242	71.7	198	32	W50252	Hepatitis B virus pre	8.45e-92
25	W50242	70.3	183	17	R68878	Hepatitis B virus E a	7.45e-90
26	W50242	70.1	183	17	R68878	Hepatitis B virus E a	1.58e-89
27	W50242	70.0	183	17	R68884	Hepatitis B virus E a	1.91e-89

RESULT 1
 ID W50242 standard: Protein; 194 AA.
 AC W50242;
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p22 polypeptide Met-p22.
 KW Hepatitis B virus precore p22 polypeptide; inhibitor; HBV; nucleocapsid; gene therapy;
 KM viral replication; liver; Met-p22.
 OS Hepatitis B virus.
 FS Synthetic.
 FH Key
 FI Protein
 FT 2.194
 PN W0909649-91.
 PD 12-MAR-1998.
 PF 03-SEP-1997; U15500.
 PR (GEO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni P, Wands JR;
 DR WPI: 98-193325/17
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 11: Page 40; 60pp: English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein with an added N-terminal Met residue. p22 is produced by
 CC elimination of the 19-amino acid leader peptide from the 25 kDa
 CC full-length HBV precore protein (see W50250). Evidence is provided
 CC that HBV replication is inhibited in the presence of high levels of
 CC HBV precore or precore-related proteins. These proteins can be
 CC incorporated into HBV nucleocapsids along with the p1 core protein
 CC (see W50251), which is the usual nucleocapsid component, and
 CC thereby render the nucleocapsids deficient in encapsidating HBV
 CC pregenomic RNA. Thus, over-expression of the precore proteins, or
 CC certain variants of them, leads to transdominant inhibition of HBV
 CC replication. Suitable inhibitory proteins include p25 (see W50250),
 CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
 CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously

CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 194 AA;

Query Match 76.6%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 6.00e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmdidpykfeafatvellsfipdsffsvrydlldtasalyrealsepncsphtalrqa1 70
 |||||||
 QY 57 SMDIDPYKFEAFATVELLSPDSFFSVRYDLDTASALYREALSEPNCSPHTALRQAI 116
 |||||||
 Db 71 lwcgelmrlatwgvnledpdsrdlvsyvnltmglkfrrqlwfhiscitfgretvley1 130
 |||||||
 QY 117 LCGEMLMTLATWGVNLEDPSRDLVSYVNTNMGKRLQMLMHISCLTGFRETVLEYL 176
 |||||||
 Db 131 vsfgvwtppayrppnapilsltpettvrrrgsprrrrrsgprrrrrsgsr 190
 |||||||
 QY 177 VSFGVWIRTPPAYRPPNAPILSTLPETTVRRRGSRPRRRRSPRRRSQSR 236
 |||||||
 Db 191 esgc 194
 |||||
 QY 237 ESQC 240

RESULT 2
 ID W50250 standard; Protein; 212 AA.
 AC W50250;

DE Hepatitis B virus precore p25 polypeptide.
 KM Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
 OS Hepatocyte; liver; p25 protein.
 KM Hepatitis B virus.
 PN W09809648-A1.
 PF 12-MAR-1998.
 PR 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI; 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 15; Page 35; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SQ Sequence 212 AA;

Query Match 76.6%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 6.00e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmdidpykfeafatvellsfipdsffsvrydlldtasalyrealsepncsphtalrqa1 88
 |||||||
 QY 57 SMDIDPYKFEAFATVELLSPDSFFSVRYDLDTASALYREALSEPNCSPHTALRQAI 116
 |||||||

Db 89 lwcgelmrlatwgvnledpdsrdlvsyvnltmglkfrrqlwfhiscitfgretvley1 148
 |||||||
 QY 117 LCGEMLMTLATWGVNLEDPSRDLVSYVNTNMGKRLQMLMHISCLTGFRETVLEYL 176
 |||||||
 Db 149 vsfgvwtppayrppnapilsltpettvrrrgsprrrrrsgprrrrrsgsr 208
 |||||||
 QY 177 VSFGVWIRTPPAYRPPNAPILSTLPETTVRRRGSRPRRRRSPRRRSQSR 236
 |||||||
 Db 209 esgc 212
 |||||
 QY 237 ESQC 240

RESULT 3
 ID R27473 standard; Protein; 346 AA.
 AC R27473;
 DT 24-FEB-1993 (first entry)
 DE S12/core protein.
 KM Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KM vaccinia virus; 13L; promoter; NYVAC; recombinant; HBV L;
 KM large pre-S antigen; spsAg; fusion protein; pre-S region; S12/core;
 KM S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KM deletion loci; recipient loci.
 OS Synthetic.
 FH Key
 FT Location/Qualifiers
 FT region 1..108
 FT /label S1
 FT region 109..163
 FT /label S2
 FT region 164..346
 FT /label Core
 PN W09215672-A.
 PD 17-SEP-1992.
 PF 09-MAR-1992; U01906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRO-) VIROGENETICS CORP.
 CO Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Lambach KJ, Norton EK, Paolletti E, Perkus ME, Pincus SE,
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI; 92-331718/40.
 DR NPSB; Q29105.

PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure; Fig 13; 45pp; English.
 CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the 13L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NYVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, spsAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NYVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC Q35501-864.
 SQ Sequence 346 AA;

Query Match 76.6%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 6.00e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 nmdidpykfeafatvellsfipdsffsvrydlldtasalyrealsepncsphtalrqa1 222


```

QY 57 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 116
DB 223 LCMGELMTLATWGVNLEDPASRDLYVSYVNTMGLKFRQLLWFHISCLTFFGREVIEWEY 282
QY 117 LCMGELMTLATWGVNLEDPASRDLYVSYVNTMGLKFRQLLWFHISCLTFFGREVIEWEYL 176
DB 283 VSFQWITPPAYRPPNAPILSTLPETVVRIRGRSPRRRTSPRRRSQSPRRRSQSR 342
QY 177 VSFQWITPPAYRPPNAPILSTLPETVVRIRGRSPRRRTSPRRRSQSPRRRSQSR 236
DB 343 ESQC 346
QY 237 ESQC 240

RESULT 4
ID P80959 standard: protein: 184 AA.
AC P80959.
DT 19-NOV-1990 (first entry)
DE Hepatitis B virus subtype ayw. core protein.
KM Conjugate: fusion protein: hepatitis B core protein; HBV subtype ayw.;
  T cell stimulating polypeptide; vaccines.
OS Synthetic.
PN EP-271302-A.
PD 15-JUN-1988.
PF 07-DEC-1987: 310725.
PR 07-OCT-1987: US-106538.
PR 07-OCT-1987: US-939617.
PA (SCRT-) Scripps Clinic Res.
PI Thornton GB, Mortuary AM, Millich DR, McLachlan A.;
DR WPI: 88-163287/24.
PT New conjugates and fusion proteins of immunogenic polypeptide -
  PT and hepatitis B core antigen and T cell stimulating polypeptide
  corep. to core antigen, useful in vaccines.
PS Disclosure: P: English.
CC This sequence contains the T cell stimulating epitopes, amino acid
  CC residue 1-55 and 70-140. It is believed that the regions 1-44 and
  CC 70-140 do not contain determinants that suppress T cell activation.
  CC Polypeptides essentially consisting of 15 to 55 amino acids
  CC corresponding to the above mentioned HBV regions are T cell
  CC stimulating. Coupling a polypeptide immunogen to such sequences,
  CC e.g. by using a bifunctional reagent which forms a disulphide link,
  CC improves its immunogenicity. These are useful in vaccines and can be used
  CC therapeutically to improve T cell response to HBcAg in infected
  CC subjects.
CC See also P80896-P80898 and P80951-P80859.
SQ Sequence 184 AA:

Query Match 75.9%; Score 1363; DB 1; Length 184;
Best Local Similarity 99.5%; Pred. No. 5, 67e-98;
Matches 183; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

DB 10 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 69
QY 57 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 116
DB 70 LCMGELMTLATWGVNLEDPASRDLYVSYVNTMGLKFRQLLWFHISCLTFFGREVIEWEY 129
QY 117 LCMGELMTLATWGVNLEDPASRDLYVSYVNTMGLKFRQLLWFHISCLTFFGREVIEWEYL 176
DB 130 VSFQWITPPAYRPPNAPILSTLPETVVRIRGRSPRRRTSPRRRSQSPRRRSQSR 189
QY 177 VSFQWITPPAYRPPNAPILSTLPETVVRIRGRSPRRRTSPRRRSQSPRRRSQSR 236
DB 190 ESQC 193
QY 237 ESQC 240

RESULT 6
ID W50251 standard: Protein: 183 AA.
AC W50251.
DT 28-SEP-1998 (first entry)
DE Hepatitis B virus p21 core protein.
KM Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
  KM hepatocyte; liver; p21; core protein.
OS Hepatitis B virus.
FT Key Location/Qualifiers
  FT 82..98
  FT /note="Immunodominant region"
  PN W09809649-AL.
  PD 12-MAR-1998.
  PF 03-SEP-1997: U15500.
  PR 03-SEP-1996: US-025370.
  PA (GEO) GEN HOSPITAL CORP.
  PI Melegari M, Scaglioni PP, Wands JR;
  DR WPI: 98-193325/17.
  PT DNA encoding proteins which can be incorporated with wild type
  PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
  PT inhibition of viral replication, especially hepatitis B virus
  PS Claim 9; Page 34-35; 60pp; English.
  CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
  CC peptide from the 25 kDa full-length HBV precore protein (see
  CC W50250). Evidence is provided that HBV replication is inhibited in
  CC the presence of high levels of HBV precore or precore-related
  CC proteins. These proteins can be incorporated into HBV nucleocapsids
  CC along with the p21 core protein (see W50251), which is the usual
  CC nucleocapsid component, and thereby render the nucleocapsids
  CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
  CC expression of the precore proteins, or certain variants of them,
  CC leads to transdominant inhibition of HBV replication. Suitable
  CC inhibitory proteins include p25 (see W50250), p22, Met-p22 (see
  CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
  CC (see W50238). Heterologous peptides (see W50244-49) may be
  CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
  CC proteins can be produced by recombinant methods using claimed
  CC expression vectors and host cells. They can be provided exogenously
  CC to the target cells for use in inhibiting HBV replication.
  CC Alternatively, a nucleic acid construct that directs overexpression
  CC of an inhibitory protein in target cells is used for the gene
  CC therapy of HBV infection.
  SQ Sequence 193 AA:

Query Match 75.7%; Score 1359; DB 32; Length 193;
Best Local Similarity 98.4%; Pred. No. 1, 20e-97;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```


PF 1-JAN-1989; 123526.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOI) Biogen Inc.
 PI Murray K. Schaller HE;
 DR WPI: 90-195067/26.
 N-PSDB: 004799
 PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
 PS used in detection of infection and in vaccine prodn.
 CC Polypeptide fragments encoded by the DNA sequence are antigenic for
 CC HBV and may be used as a vaccine or in detection. Peptides may be
 CC cultured in a suitable bacterial host such as E.coli.
 CC Fragments of the sequence are also claimed as being antigenically
 CC useful.
 SQ Sequence 183 AA;

Query Match 75.1%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 7.78e-97;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdldpykefatvellsflpsdffsvrrdlldaalyrdlaespehscphthalrqa11 60
 |||
 QY 58 MDIDPYKEFATVLLSFLPSDFPSVRDLDTASALYREALSPERHSCPHHTALRQA1L 117
 Db 61 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 120
 |||
 QY 118 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 177
 Db 121 sfgywrtppayrpnaplslpctvrrrgsprrrtpprrrrsgsrr 180
 |||
 QY 178 SFGWIRTPPAYRPPNAPILSTPETTVRRGRSPRRRTPSPRRRSQSRRSRE 237
 Db 181 sqc 183
 |||
 QY 238 SQC 240

RESULT 9
 ID P00041 standard; Protein; 183 AA.
 AC P00041;
 DT 14-OCT-1992 (first entry)
 DE Sequence of core antigen.
 KM Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
 OS Hepatitis B virus.
 PS EP-13828-A.
 PN 06-AUG-1980.
 PR 21-DEC-1979; 303017.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOI) BIOGEN NV.
 PI Murray K. Schaller HE;
 DR WPI: 80-57268C/33.
 N-PSDB: N00003.
 PT Recombinant DNA coding for polypeptide - have specificity of
 PT hepatitis B viral antigens in detection or antibody stimulation
 PS Example: Figs 3-4; 43pp; English.
 CC Human serum from a single HBsAg positive, HBsAg positive donor
 CC (serotype adym) was used to prep. a DNA-contg. pellet which was
 CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
 CC The labelled DNA was then extracted with phenol from the resulting
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
 CC It was then cloned in plasmid pBR322 which was used to transform E.
 CC coli. Micro-organisms prepd. by the processes are deposited at the
 CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg:
 CC HBV-Kpn I dg: Tetr Amps HBV+.
 SQ Sequence 183 AA;

Query Match 75.1%; Score 1349; DB 5; Length 183;
 Best Local Similarity 96.7%; Pred. No. 7.78e-97;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdldpykefatvellsflpsdffsvrrdlldaalyrdlaespehscphthalrqa11 60
 |||
 QY 58 MDIDPYKEFATVLLSFLPSDFPSVRDLDTASALYREALSPERHSCPHHTALRQA1L 117
 Db 61 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 120
 |||
 QY 118 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 177
 Db 121 sfgywrtppayrpnaplslpctvrrrgsprrrtpprrrrsgsrr 180
 |||
 QY 178 SFGWIRTPPAYRPPNAPILSTPETTVRRGRSPRRRTPSPRRRSQSRRSRE 237
 Db 181 sqc 183
 |||
 QY 238 SQC 240

RESULT 10
 ID W09048 standard; Protein; 397 AA.
 AC W09048;
 DT 11-APR-1997 (first entry)
 DE Plasmid pHBV DN AA encoded HBV dominant negative polypeptide.
 KM Hepadnavirus; woodchuck hepatitis virus; hepatitis delta virus;
 KM HBV; ground squirrel hepatitis B virus; duck hepatitis B virus;
 KM core protein; replication; antiviral; gene therapy; pHBV DN AA.
 OS Hepatitis B virus.
 FH Key
 FT location/Qualifiers
 FT region 1..179
 FT /label="HBV-core
 FT /note="positions 1-179 correspond to amino acids
 FT 1-179 of HBV core protein"
 FT region 180..397
 FT /note="positions 180-397 correspond to amino
 FT acids 9-226 of HBV surface protein"

PN W09700698-A1.
 PD 09-JAN-1997.
 PF 20-JUN-1996; U10602.
 PR 20-JUN-1995; US-017814.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 97-087176/08.
 N-PSDB: T49598.
 PT New method for inhibiting the replication of hepadnaviruses -
 PT comprises introducing a mutant polypeptide with a mutated core
 PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
 PT B
 PS Disclosure; Page 46-48; 83pp; English.
 CC Plasmid pHBV DN AA carries an insert (T49598) coding for a protein
 CC (W09048) comprising a hepatitis B virus (HBV) core protein (see
 CC also W09044) fused in-frame at amino acid 179 with the HBV surface
 CC protein (see also W09045) at amino acid 8. Plasmid pHBV DN BB
 CC (T49599) expresses the HBV core fused at amino acid 175 to the
 CC surface protein at amino acid 112 (W09049). pHBV DN AA was at
 CC least as potent an inhibitor of HBV replication as construct
 CC pHBV DN (T49597); pHBV DN BB was less inhibitory than pHBV DN.
 CC Vectors expressing hepadnavirus dominant negative core mutants can
 CC be utilised in the gene therapy of viral infections.
 SQ Sequence 397 AA;

Query Match 74.8%; Score 1344; DB 20; Length 397;
 Best Local Similarity 100.0%; Pred. No. 1.98e-96;
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdldpykefatvellsflpsdffsvrrdlldaalyrdlaespehscphthalrqa11 60
 |||
 QY 58 MDIDPYKEFATVLLSFLPSDFPSVRDLDTASALYREALSPERHSCPHHTALRQA1L 117
 Db 61 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 120
 |||
 QY 118 CWGELMTLATWGVNLDPSRDLVSYVNTNMGKROLMFHISCLTGREVIEYLV 177
 Db 121 sfgywrtppayrpnaplslpctvrrrgsprrrtpprrrrsgsrr 179

Query 178 SFGWIRTPAPRPAPNAPILSTLPEITVVRKGRSPRRRTSPRRRSOSPRRRRSOSR 236

RESULT 11
ID P00004 standard; Protein: 184 AA.

AC P00004;
DT 14-OCT-1992 (first entry)
DE Sequence of core antigen.
KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
OS Hepatitis B virus.

PN EP-13828-A.
PD 06-AUG-1980.
PF 21-DEC-1979; 303017.
PR 22-DEC-1978; GB-049907.
PR 27-DEC-1978; GB-050039.
PR 01-NOV-1979; GB-037910.
PA (BIOJ) BIOGEN NV.
PI Murray K, Schaller HE.
DR WPI: 80-57268C/33.
DR N-PSDB: N00002.
PT Recombinant DNA coding for polypeptide - have specificity of
PT hepatitis B viral antigens in detection or antibody stimulation
PS Claim 13: Page 40; 43pp: English.

CC Human serum from a single HBsAg positive, HBeAg positive donor
CC (serotype adym) was used to prep. a DNA-contg. pellet which was
CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
CC The labelled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prepd. by the processes are deposited at the
CC NCIB as pBR322-HBV-G⁺-L, e.g. E. coli HB101/pBR322-Pst I dg:
CC HBV-Kpn I dc: TetA Amps HBV+.
SQ Sequence 184 AA;

Query Match 74.5%; Score 1338; DB 5; Length 184;

Best Local Similarity 96.2%; Pred. No. 6.09e-96;
Matches 177; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

DB 1 mdidpykefgatvellslfspdffpsvrcdlldtaaalysdalespeshphthalrgai 60
QY 58 MDIDPYKEFGATVELLSFLPSDFPSV-RDLDTASALYREALSEPHCHPHHTALRQAI 116
DB 61 lcgwlmnlatawgsnledpasrdlvsyvnmgklrqlwfhiscitfgretvleyl 120
QY 117 LCGELMTLATWGVNLDPASRDLVSYVNTMGLKFRQLMFHISCITFGRETVEYL 176
DB 121 vsfgwlrtpaprpapnapijstlpetvvrgrgsprrrtsprrrrsgsr 180
QY 177 VSGWLRTPAPRPAPNAPILSTLPEITVVRKGRSPRRRTSPRRRSOSPRRRRSOSR 236
DB 181 esqc 184
QY 237 ESQC 240

RESULT 12
ID R40806 standard; Protein: 196 AA.

AC R40806;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 fusion.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FT Key Location/Qualifiers
FT region 7..13
FT protein /label= PV-1
FT protein /label= HBC

PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISM) NISSHIN OIL MILLS LTD. ?

DR WPI: 93-277479/35.

DR N-PSDB: Q47736.
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure: Fig 7; 12pp: Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.
SQ Sequence 196 AA;

Query Match 74.4%; Score 1336; DB 8; Length 196;
Best Local Similarity 95.1%; Pred. No. 8.86e-96;
Matches 175; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

DB 13 andidpykefgasvellslfspdffpsvrcdlldtaaalysdalespeshphthalrgai 72
QY 57 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCHPHHTALRQAI 116
DB 73 lcgwlmnlatawgsnledpasrdlvsyvnmgklrqlwfhiscitfgretvleyl 132
QY 117 LCGELMTLATWGVNLDPASRDLVSYVNTMGLKFRQLMFHISCITFGRETVEYL 176
DB 133 vsfgwlrtpaprpapnapijstlpetvvrgrgsprrrtsprrrrsgsr 192
QY 177 VSGWLRTPAPRPAPNAPILSTLPEITVVRKGRSPRRRTSPRRRSOSPRRRRSOSR 236
DB 193 esqc 196
QY 237 ESQC 240

RESULT 13
ID R40808 standard; Protein: 208 AA.

AC R40808;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 / IL-1 fusion.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FT Key Location/Qualifiers
FT region 8..19
FT region /label= PV-1
FT region 20..25
FT protein /label= IL-1
FT protein 26..208
FT protein /label= HBC

PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISM) NISSHIN OIL MILLS LTD.
DR WPI: 93-277479/35.

DR N-PSDB: Q47738.
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure: Fig 9; 12pp: Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.
SQ Sequence 208 AA;

Query Match 74.4%; Score 1337; DB 8; Length 208;
Best Local Similarity 94.6%; Pred. No. 7.35e-96;
Matches 175; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

DB 24 imdidpykefgasvellslfspdffpsvrcdlldtaaalysdalespeshphthalrgai 83
QY 56 LMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCHPHHTALRQAI 115
DB 84 lcgwlmnlatawgsnledpasrdlvsyvnmgklrqlwfhiscitfgretvleyl 143

[illegible]

RESULT 14 standard: Protein; 183 AA.
ID P40311;
AC P40311;
DT 18-AUG-1992 (first entry)
DE Hepatitis virus core antigen.
KW HBcAg; vaccine; diagnosis; HBV infection.
OS Hepatitis b virus.
PN J58074985-A.
PD 27-APR-1984.
PF 19-OCT-1982; 183432.
PR 19-OCT-1982; JP-183432.
PA (TAKE) TAKEEDA CHEMICAL IND KK.
DR WPI: 84-143231/23.
PT DNA used in prevention of infections by hepatitis virus B -
PT comprises structural gene of hepatitis virus adr B surface antigen
PT coding gene and at least 1 virus core antigen structural gene.
PS Disclosure: Fig 3; 13pp; Japanese.
CC The sequence is that of hepatitis virus core antigen (HBcAg) . It
CC can be used as a vaccine for the prevention of infections by
CC hepatitis B virus (HBV) and also in the diagnosis of early stages
CC of HBV infection. See also P40310.
QC Sequence 183 AA;

Query Match	74.38;	Score 1334;	DB 4;	Length 183;
Best Local Similarity	95.68;	Pred. No. 1.29e-95;		
Matches 175;	Conservative 5;	Mismatches 3;	Indels 0;	Gaps 0;

[illegible]

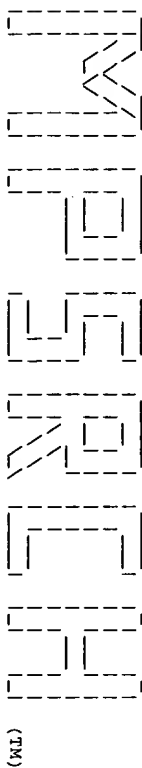
RESULT	15	standard: Protein: 183 AA.
ID	RA0805	
AC	RA0805;	
DI	16-FEB-1994	(first entry)
DE	Hepatitis B core protein.	
KW	Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.	
OS	Hepatitis B virus.	
PN	J05192170-A.	
PD	03-AUG-1993.	
PF	24-SEP-1991;	243800.
PR	24-SEP-1991;	JP-243800.
PA	(NISW)	NISSHIN OIL MILLS LTD.
DR	WPI;	93-277479/35.
DR	N-PSDB;	Q47735.
PT	Recombinant plasmid for high immunogenicity virus - contains	
PT	recombined haemagglutinin gene, hepatitis B core gene of vaccinia	
PT	virus and exocytic genes	

PS Disclosure: Fig 6: 12pp: Japanese.
 CC The hepatitis B core gene is recombined with pV-1 DNA
 CC and IL-1 beta to form a plasmid (047738) which is then used to produce
 CC a vaccine. NB: Sequence is difficult to read in the original
 CC specification.
 50 Sequence 183 AA;

Query Match	74.38;	Score 1334;	DB 8;	Length 183;
Best Local Similarity	95.68;	Pred. No. 1.29e-95;		
Matches 175; Conservative	5;	Mismatches 3;	Indels 0;	Gaps 0;

Dd		1	mdidpyvefagssvallisflsfodffosridlldesalryealspepcspbhcalqaii	60
Oy		58	MDIDPYVEFGATVLSLFSLFSDFFPSRDLLDTRASALRYEALSPEPCSPBHTRALKQAIL	117
Dd		61	cwgeimlatwvsnglsedpsarelvsvayvvvmgklkrlqlfwlfscitfgretyleiy	120
Oy		118	CWGEIMLATATWGVNLEDDPASRDLVSVYVMTNMGLKROLFWHISCLTFRGRETYLEIYV	177
Dd		121	sfgvwlrtpypayrzpnnapllstlbectevrrirgrsprrrtpprrrsgpprrrrsgare	180
Oy		178	SFGVWLRTPPAYRZPNAPPIILSTLBETIVVRRRGRSPRRRTPPRRRRSOSPRRRRSSORE	237
Dd		181	sqc 183	
Oy		238	SQC 240	

Search completed: Thu Dec 16 13:11:36 1999
Job time : 162 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_p protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:11:54 1999; Maspar time 12.41 Seconds
775.040 Million cell updates/sec
Tabular output not generated.

Title: >US3835-10-38183
Description: (1-240) from us3835-10-38183.pep
Perfect Score: 1796
Sequence: 1 MDIDPKKFGATVELLSFLP.....RRRSQSPRRRSQSRSESQC 240

Scoring table:
PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r60
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 47.190; Variance 119.785; scale 0.394

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1375	76.6	212	1	NKVLAH e antigen precursor / 4.47e-198	
2	1374	76.5	212	2	S53211 e antigen precursor / 6.55e-198	
3	1374	76.5	212	2	S53204 e antigen precursor / 6.55e-198	
4	1373	76.4	183	2	S53207 core antigen - hepati / 9.59e-198	
5	1373	76.4	212	2	S520750 e antigen precursor / 3.01e-197	
6	1370	76.3	183	2	S53216 e antigen precursor / 3.01e-197	
7	1370	76.3	212	2	S53216 e antigen precursor / 3.01e-197	
8	1370	76.3	212	2	S53216 e antigen precursor / 3.01e-197	
9	1369	76.2	212	2	S53163 e antigen precursor / 4.40e-197	
10	1369	76.2	212	2	S520746 e antigen precursor / 4.40e-197	
11	1369	76.2	212	2	S53225 e antigen precursor / 4.40e-197	
12	1368	76.2	212	2	S53198 e antigen precursor / 6.45e-197	
13	1366	76.1	212	2	S53223 e antigen precursor / 1.38e-196	
14	1366	76.1	212	2	S53223 e antigen precursor / 1.38e-196	
15	1364	75.9	183	2	S53169 core antigen - hepati / 2.96e-196	
16	1363	75.9	183	2	S53247 core antigen - hepati / 4.33e-196	
17	1363	75.9	212	2	S53202 e antigen precursor / 4.33e-196	
18	1363	75.9	212	2	S53159 e antigen precursor / 4.33e-196	
19	1363	75.9	212	2	S53200 e antigen precursor / 4.33e-196	
20	1362	75.8	212	2	S53251 e antigen precursor / 6.34e-196	
21	1360	75.7	212	2	S53204 e antigen precursor / 1.36e-195	
22	1360	75.7	212	2	S53227 e antigen precursor / 1.36e-195	
23	1359	75.7	212	2	S53233 e antigen precursor / 1.99e-195	

24	1358	75.6	211	1	NKVLAI	e antigen precursor	/	2.91e-195
25	1357	75.6	212	2	S53281	e antigen precursor	/	4.26e-195
26	1355	75.4	212	1	NKVLBI	e antigen precursor	/	9.13e-195
27	1355	75.4	212	2	S53242	e antigen precursor	/	9.13e-195
28	1351	75.2	183	2	S53129	core antigen - hepati	/	4.19e-194
29	1351	75.2	212	2	S53240	e antigen precursor	/	4.19e-194
30	1349	75.1	183	1	NKVLAI2	core antigen - hepati	/	8.98e-194
31	1349	75.1	183	2	S53181	core antigen - hepati	/	8.98e-194
32	1348	75.1	183	2	S53232	core antigen - hepati	/	1.31e-193
33	1348	75.1	183	2	S53260	core antigen - hepati	/	1.31e-193
34	1345	74.9	183	2	S53152	core antigen - hepati	/	4.12e-193
35	1343	74.8	183	2	S53214	core antigen - hepati	/	8.83e-193
36	1344	74.8	212	2	S53229	e antigen precursor	/	6.03e-193
37	1342	74.7	183	2	S53140	core antigen - hepati	/	1.29e-192
38	1341	74.7	183	2	S53267	core antigen - hepati	/	1.89e-192
39	1341	74.7	183	2	S53189	core antigen - hepati	/	1.89e-192
40	1341	74.7	212	2	S53238	e antigen precursor	/	1.89e-192
41	1339	74.6	183	2	S53137	core antigen - hepati	/	4.05e-192
42	1340	74.6	212	2	S53257	e antigen precursor	/	2.77e-192
43	1339	74.6	212	2	S25651	e antigen precursor	/	4.05e-192
44	1338	74.5	212	2	S53236	e antigen precursor	/	5.93e-192
45	1337	74.4	212	1	NKVLJ2	e antigen precursor	/	8.68e-192

REFERENCE S47404 Plucieniczak, A.
#authors submitted to the EMBL Data Library, August 1994
#description Molecular cloning and sequencing of two complete genomes of
#accession S47405
##molecule_type DNA
##molecule_type DNA
##residues 1-212 ##label PLU
##cross-references EMBL:235116; NID:q527435; PID:q527437
##experimental_source subtype ayw4, isolate hb321
REFERENCE S53112
#authors Lal, M.E.; Mazzoleni, A.P.; Porru, A.; Baleshtierl, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53191
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85290; NID:q736088; PID:q736090
##experimental_source isolate patient Ferracuti/83
#accession S53209
##molecule_type DNA
##residues 1-212 ##label LA2
##cross-references EMBL:X85290; NID:q736114; PID:q736116
##experimental_source isolate patient Castag/83
#accession S53234
##molecule_type DNA
##residues 1-212 ##label LA3
##cross-references EMBL:X85300; NID:q736150; PID:q736152
##experimental_source isolate patient Sanna/84
#accession S53264
##molecule_type DNA
##residues 1-212 ##label LA4
##cross-references EMBL:X85313; NID:q736194; PID:q736196

```
##experimental_source isolate patient Licheri-1'85
#accession S53249
##molecule_type DNA
##residues 1-212 ##label LA5
##cross-references EMBL:X85306; NID:g736172; PID:g736174
##experimental_source isolate patient Flore-1'86
#accession S53262
##molecule_type DNA
##residues 1-212 ##label LA6
##cross-references EMBL:X85312; NID:g736191; PID:g736193
##experimental_source isolate patient Licheri'83
#accession S53277
##molecule_type DNA
##residues 30-212 ##label LA7
##cross-references EMBL:X85317; NID:g736211; PID:g736214
##experimental_source patient Giordo-2'86
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be processed

REFERENCE A93214
#authors Galibert, F.; Mandart, E.; Fitoussi, F.; Tellois, P.;
#journal Charney, P.
#title Nature (1979) 281:646-650
#feature Nucleotide sequence of the hepatitis B virus genome (subtype
#cross-references M01D:81012091
#accession A03711
##molecule_type DNA
##residues 1-212 ##label GAL
##cross-references GB:J02203; NID:g329640; PID:g329642
#experimental_source subtype ayw

GENETICS
#gene C
#classification #superfamily hepatitis B virus core antigen
#keywords alternative initiators; core protein
FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label EAG\
#domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular-weight 24350 #checksum 782

Query Match 76.6%; Score 1375; DB 1; Length 212;
Best Local Similarity 99.5%; Pred. No. 4.47e-198;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKFGATVELLFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
:|||||
QY 57 SMIDPKKKGATVELLFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 116
|||||
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVEYL 148
|||||
QY 117 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVEYL 176
|||||
Db 149 VSPGVWIRPPAPRPNAPILSTLPTTVVRRGRSPRRRPSRRRSOSPRRRSOSR 208
|||||
QY 177 VSPGVWIRPPAPRPNAPILSTLPTTVVRRGRSPRRRPSRRRSOSPRRRSOSR 236
|||||
Db 209 ESOC 212
|
QY 237 ESOC 240

RESULT 2
ENTRY S53211 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Castag-1'85 and others)
ALTERNATE_NAMES HBC antigen; HBe antigen precursor / HBC antigen; pre-C/C
antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castag-1'85; isolate patient Ferracuti-1'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
```

```
08-Sep-1997
S53211; S53197
ACCESSIONS
REFERENCE
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53211
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85291; NID:g736117; PID:g736119
##experimental_source isolate patient Castag-1'85
#accession S53197
##molecule_type DNA
##residues 30-212 ##label LAW
##cross-references EMBL:X85284; NID:g736095; PID:g736098
##experimental_source isolate patient Ferracuti-1'89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
#classification #superfamily hepatitis B virus core antigen
#keywords alternative initiators; core protein
FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label EAG\
#domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular-weight 24336 #checksum 746

Query Match 76.5%; Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 6.55e-198;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKFGATVELLFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
:|||||
QY 57 SMIDPKKKGATVELLFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 116
|||||
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVEYL 148
|||||
QY 117 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVEYL 176
|||||
Db 149 VSPGVWIRPPAPRPNAPILSTLPTTVVRRGRSPRRRPSRRRSOSPRRRSOSR 208
|||||
QY 177 VSPGVWIRPPAPRPNAPILSTLPTTVVRRGRSPRRRPSRRRSOSPRRRSOSR 236
|||||
Db 209 ESOC 212
|
QY 237 ESOC 240

RESULT 3
ENTRY S32204 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1000)
ALTERNATE_NAMES HBe antigen precursor / HBC antigen; pre-C/C antigen
core antigen; e antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient C1000
#subtype ayw, patient C1000
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997
S32204
ACCESSIONS
REFERENCE
#authors Preisler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.;
Gerok, W.; Rasenack, J.
#submission submitted to the EMBL Data Library, March 1993
#accession identified and sequence analysis of hepatitis B virus DNA
in immunological negative infection.
#accession S32204
##molecule_type DNA
##residues 1-212 ##label PRE
##cross-references EMBL:X72702; NID:g288927; PID:g288930
##experimental_source subtype ayw, patient C1000
#note due to a stop codon between the alternative initiators
```


GENETICS the e antigen precursor cannot be produced

CLASSIFICATION C
#gene #superfamily hepatitis B virus core antigen
KEYWORDS core protein
FEATURE 1-29

30-212 #domain signal sequence #status predicted #label sig\
30-178 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link ENG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24336 #checksum 770

Query Match 76.5%; Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 6.55e-198;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVELLSFLPSDFPVSVDLDTASALYDALESPEHCSPHHTALROAI 88
117 LCGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 176
149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
177 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 236

DB 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 148
117 LCGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 176
149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
177 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 236
DB 209 ESOC 212
111
QY 237 ESOC 240

RESULT 4
ENTRY S53207 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
Castaa-2'87)

ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Castaa-2'87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53207
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Portu, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53207

##molecule_type DNA
##residues 1-183 ##label LAI
##cross-references EMBL:X85289; NID:G736110; PID:G736113
##experimental_source isolate patient Castaa-2'87
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular-weight 21102 #checksum 2165

Query Match 76.4%; Score 1373; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 9.59e-198;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPKFEGATVELLSFLPSDFPVSVDLDTASALYDALESPEHCSPHHTALROAI 60
58 MDIDPKFEGATVELLSFLPSDFPVSVDLDTASALYDALESPEHCSPHHTALROAI 117
DB 61 CNGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 120
QY 118 CNGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 177

DB 121 SFGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 180
QY 178 SFGWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSRE 237
DB 181 SOC 183
QY 238 SOC 240

RESULT 5
ENTRY S20750 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1)
ALTERNATE_NAMES HBE antigen precursor / HBC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient C1
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997

ACCESSIONS S20750
REFERENCE S20745
#authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.;
Portu, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with
HBsAg negative chronic liver disease.
#accession S20750
##molecule_type DNA
##residues 1-212 ##label LAI
##cross-references EMBL:X5258; NID:G59434; PID:G59436
##experimental_source subtype ayw, patient C1

GENETICS C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE 1-29
30-212 #domain signal sequence #status predicted #label sig\
30-178 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link ENG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 147
Query Match 76.4%; Score 1373; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 9.59e-198;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVELLSFLPSDFPVSVDLDTASALYDALESPEHCSPHHTALROAI 88
57 SMDIDPKFEGATVELLSFLPSDFPVSVDLDTASALYDALESPEHCSPHHTALROAI 116
DB 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 148
QY 117 LCGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLTFGRETVEYL 176

DB 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 177 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 236
DB 209 ESOC 212
111
QY 237 ESOC 240

RESULT 6
ENTRY S53270 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
Licheri-2'87)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Licheri-2'87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

```

ACCESSIONS      08-Sep-1997
REFERENCE       S533270
#authors       S53112
#submitter     Lai, M.E.; Mazzoletti, A.P.; Porru, A.; Balestrieri, A.
#description   Submitted to the EMBL Data Library, March 1995
#accession     S53270
##molecule_type DNA
#residues      1-183 #label LAI
##cross-references EMBL:X85314; NID:g736201; PID:g736204
##experimental_source isolate patient Licheri-2/87
#note          due to a stop codon between the alternative initiators
               the e antigen precursor cannot be produced

GENETICS
#gene          C
CLASSIFICATION #superfamily hepatitis B virus core antigen
               core protein
KEYWORDS       #length 183 #molecular-weight 21102 #checksum 2199
SUMMARY

Query Match      76.3%; Score 1370; DB 2; Length 183;
Best Local Similarity 98.4%; Pred. No. 3,01e-197;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0

Db 1 MDIDYKEFGATVELLSLPDFFPSVDLIDTASALYRDALSPHCSPHNTALROAIL 60
Oy 58 MDIDYKEFGATVELLSLPDFFPSVDLIDTASALYRDALSPHCSPHNTALROAIL 117
    |||
Db 61 CWMGLMTATATGAVGLDEPPASRDLYVSVNTNMGLKFKQLLMFHSICLTPEREVIEYLV 120
Oy 118 CWMGLMTATATGAVGLDEPPASRDLYVSVNTNMGLKFKQLLMFHSICLTPEREVIEYLV 177
    |||
Db 121 SFGVWIRTPPAPRPANAILSTLETTYVRRRGRTPRRTPSPRRRSQSPRRRSQSRE 180
Oy 178 SFGVWIRTPPAPRPANAILSTLETTYVRRRGSRPRRTPSPRRRSQSPRRRSQSRE 237
    |||
Db 181 SOC 183
Oy 238 SOC 240

RESULT 7
ENTRY   S53216 #type complete
TITLE   e antigen precursor / core antigen - hepatitis B virus
        (isolate patient Castag'3)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen; pre-C/C antigen
CONTAINS      core antigen; e antigen
ORGANISM      #formal_name hepatitis B virus; HBV
               #variety isolate patient Castag'3
DATE          08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
               08-Sep-1997
ACCESSIONS    S53216
REFERENCE     S53112
#authors      Lai, M.E.; Mazzoletti, A.P.; Porru, A.; Balestrieri, A.
#submitter    Submitted to the EMBL Data Library, March 1995
#accession    S53216
##molecule_type DNA
#residues     1-212 #label LAI
##cross-references EMBL:X85293; NID:g736124; PID:g736126
##experimental_source isolate patient Castag'3

CLASSIFICATION C
               #superfamily hepatitis B virus core antigen
               alternative initiators; core protein
KEYWORDS
FEATURE
1-29          #domain signal sequence #status predicted #label SIG\
30-212        #product core antigen #status predicted #label CAG\
30-178        #product e antigen #status predicted #label EAG\
179-212       #domain carboxyl-terminal propeptide #link EAG #status
               predicted #label ECP
SUMMARY       #length 212 #molecular-weight 24363 #checksum 752

Query Match      76.3%; Score 1370; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 3,01e-197;
Matches 182; Conservative 1; Indels 0; Gaps 0

```

Db	29	GMDIDPYKEFGATVLLSFLPSDFSPSRDLIDLNASALYREALSEPCSPHNTLRQAI	88
QY	57	SMDIDPYKEFGATVLLSFLPSDFSPSRDLIDLNASALYREALSEPCSPHNTLRQAI	116
Db	89	LCWGLMLTLATWGVNLEDPSASRLVSYVNTNMGLKFRQLMFIISCLTGRETVIEYL	148
QY	117	LCWGLMLTLATWGVNLEDPSASRLVSYVNTNMGLKFRQLMFIISCLTGRETVIEYL	176
Db	149	VSFQWINTPPAYRPNPAPILSTLPEITVVRRRGSSPPRRRPPSPRRRSQSPRRRSQSR	208
QY	177	VSFQWINTPPAYRPNPAPILSTLPEITVVRRRGSSPPRRRPPSPRRRSQSPRRRSQSR	236
Db	209	ESQC 212	
QY	237	ESQC 240	
RESULT	8		
ENTRY		S53272	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus (isolate patient Lichneri-3'90)	
ALTERNATE_NAMES		HBe antigen precursor / HBeC antigen; pre-C/C antigen	
CONTAINS		core antigen; e antigen	
ORGANISM		Hepatitis B virus, HBV	
VARIETY		isolate patient Lichneri-3'90	
DATE		08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		S53272	
REFERENCE		S53112	
AUTHORS		Lai, M.-E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.	
SUBMISSION		submitted to the EMBL Data Library, March 1995	
ACCESSION		S53272	
MOLECULE_TYPE		DNA	
RESIDUES		1-212	#label LAI
CROSS-REFERENCES		EMBL:X85315; NID:g736205; PID:g736207	
EXPERIMENTAL_SOURCE		isolate patient Lichneri-3'90	
GENETICS			
CLASSIFICATION		C	
KEYWORDS		#superfamily hepatitis B virus core antigen	
FEATURE		alternative initiators; core protein	
	1-29	#domain signal sequence #status predicted #label SIG\	
	30-212	#product core antigen #status predicted #label CAG\	
	30-178	#product e antigen #status predicted #label EAG\	
	179-212	#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY		#length 212	#molecular-weight 24366 #checksum 446
Query Match		76.3%	Score 1370; DB 2; Length 212;
Best Local Similarity		98.9%	Pred. No. 3, 01e-197;
Matches 182; Conservative		1; Mismatches 1; Indels 0; Gaps 0	
Db	29	GMDIDPYKEFGATVLLSFLPSDFSPSRDLIDLNASALYREALSEPCSPHNTLRQAI	88
QY	57	SMDIDPYKEFGATVLLSFLPSDFSPSRDLIDLNASALYREALSEPCSPHNTLRQAI	116
Db	89	LCWGLMLTLATWGVNLEDPSASRLVSYVNTNMGLKFRQLMFIISCLTGRETVIEYL	148
QY	117	LCWGLMLTLATWGVNLEDPSASRLVSYVNTNMGLKFRQLMFIISCLTGRETVIEYL	176
Db	149	VSFQWINTPPAYRPNPAPILSTLPEITVVRRRGSSPPRRRPPSPRRRSQSPRRRSQSR	208
QY	177	VSFQWINTPPAYRPNPAPILSTLPEITVVRRRGSSPPRRRPPSPRRRSQSPRRRSQSR	236
Db	209	ESQC 212	
QY	237	ESQC 240	
RESULT	9		
ENTRY		S53163	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus	

(isolate patient Vittorina '92)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Vittorina '92
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53163
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53163
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85256; NID:q736050; PID:q736052
#experimental_source isolate patient Vittorina '92

GENETICS C
#gene #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE 1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24322 #checksum 842

Query Match 76.2%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 4,40e-197;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
:|||||
Oy 57 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 116
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 148
:|||||
Oy 117 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 176
:|||||

Db 149 VSGFWITPPAYRPPNPIILSTPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
Oy 177 VSGFWITPPAYRPPNPIILSTPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 236
:|||||

Db 209 ESOC 212
:|||||
Oy 237 ESOC 240

RESULT 10
ENTRY S20746 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
CLASSIFICATION C
KEYWORDS (subtype ayw, patient C)
HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient C
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997

ACCESSIONS S20746
REFERENCE S20745
#authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic liver disease.
#accession S20746
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X65257; NID:q59429; PID:q59431
#experimental_source subtype ayw, patient C

GENETICS C
#gene

CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE 1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 506

Query Match 76.2%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4,40e-197;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
:|||||
Oy 57 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 116
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 148
:|||||
Oy 117 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 176
:|||||

Db 149 VSGFWITPPAYRPPNPIILSTPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
Oy 177 VSGFWITPPAYRPPNPIILSTPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 236
:|||||

Db 209 ESOC 212
:|||||
Oy 237 ESOC 240

RESULT 11
ENTRY S53225 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
CLASSIFICATION C
KEYWORDS (isolate patient Chighine '2/86)
HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Chighine-2/86
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53225
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53225
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85296; NID:q736137; PID:q736139
#experimental_source isolate patient Chighine-2/86

GENETICS C
#gene
KEYWORDS #superfamily hepatitis B virus core antigen
alternative initiators; core protein
FEATURE 1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24308 #checksum 655

Query Match 76.2%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4,40e-197;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
:|||||
Oy 57 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 116
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 148
:|||||
Oy 117 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFWHISCLTFGRETIVIEL 176
:|||||

Db 149 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 208
177 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 236
QY 209 ESOC 212
237 ESOC 240

RESULT 12
ENTRY S53198 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Ferracuti-2/90)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen: e antigen
ORGANISM #formal_name hepatitis B virus, HBV
DATE isolate patient Ferracuti-2/90
08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS
REFERENCE S53198
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53198
#molecule_type DNA
#residues 1-212 ##label LAI
#cross-references EMBL:X85285; NID:g736099; PID:g736100
#experimental_source isolate patient Ferracuti-2/90

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24292 #checksum 593

Query Match 76.2%; Score 1368; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 6,45e-197;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 88
57 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 116
QY 89 LCGDMLSLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEYL 148
117 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEYL 176
Db 149 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 208
177 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 236
QY 209 ESOC 212
237 ESOC 240

RESULT 13
ENTRY S53223 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Chighine-1/85)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen: e antigen
ORGANISM #formal_name hepatitis B virus, HBV
DATE isolate patient Chighine-1/85
08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS
REFERENCE S53223

REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53223
#molecule_type DNA
#residues 1-212 ##label LAI
#cross-references EMBL:X85295; NID:g736134; PID:g736136
#experimental_source isolate patient Chighine-1/85

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24364 #checksum 1123

Query Match 76.1%; Score 1366; DB 2; Length 212;
Best Local Similarity 97.8%; Pred. No. 1.38e-196;
Matches 180; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 88
57 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 116
QY 89 LCGDMLSLATWGVNLEDPISRDVSVYNTNMGKLFROLMEHISCLTFGRETVEYL 148
117 LCGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMEHISCLTFGRETVEYL 176
Db 149 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 208
177 VSEGVWIRTPPAYRPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQS 236
QY 209 ESOC 212
237 ESOC 240

RESULT 14
ENTRY S53274 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Giordo/84)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
CONTAINS core antigen: e antigen
ORGANISM #formal_name hepatitis B virus, HBV
DATE isolate patient Giordo/84
08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS
REFERENCE S53274
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53274
#molecule_type DNA
#residues 1-212 ##label LAI
#cross-references EMBL:X85316; NID:g736208; PID:g736210
#experimental_source isolate patient Giordo/84

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
30-178 #product e antigen #status predicted #label EAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24376 #checksum 390

Query Match 76.1%; Score 1366; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 1.38e-196;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSLPSDFPSVVDLDTASALYREALSPHHTALROAI 88
:|||||
Qy 57 SMDIDPYKEFGATVLLSLPSDFPSVVDLDTASALYREALSPHHTALROAI 116
:|||||
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
:|||||
Qy 117 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 176
:|||||
Db 149 VSEGVWIRTPPARPPNAPILSLPETVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
Qy 177 VSEGVWIRTPPARPPNAPILSLPETVVRGRGSPRRRTSPRRRSQSPRRRSQSR 236
:|||||
Db 209 ESOC 212
:|||||
Qy 237 ESOC 240

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu'89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Muresu'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53169
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
##molecule_type DNA
##residues 1-183 ##label LAI
##cross-references EMBL:X85275; NID:g736057; PID:g736060
##experimental_source isolate patient Muresu'89
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular-weight 21088 #checksum 1849

Query Match 75.9%; Score 1364; DB 2: Length 183;
Best Local Similarity 98.4%; Pred. No. 2.96e-196;
Matches 180; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDYKEFGATVLLSLPSDFPSVVDLDTASALYREALSPHHTALROAIL 60
:|||||
Qy 58 MDIDYKEFGATVLLSLPSDFPSVVDLDTASALYREALSPHHTALROAIL 117
:|||||
Db 61 CWMELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 120
:|||||
Qy 118 CWMELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 177
:|||||
Db 121 SFGVWIRTPPARPPNAPILSLPETVVRGRGTPRRRTSPRRRSQSPRRRSQSR 180
:|||||
Qy 178 SFGVWIRTPPARPPNAPILSLPETVVRGRGTPRRRTSPRRRSQSPRRRSQSR 237
:|||||
Db 181 SOC 183
:|||||
Qy 238 SOC 240

Search completed: Thu Dec 16 13:13:06 1999
Job time : 72 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

Title: >US835-10-38183
Description: (1-240) from us835-10-38183.pep
Perfect Score: 1796
Sequence: I MDIDPYEFEGATVELLSFLP.....RRRSQSPRRRRSQSRRESQC 240

```

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

```
Post-processing: Minimum Match 0%
Listing first 45 summaries
```

```
Database: swiss-prot37
1:swissprot
```

Statistics: Mean 48.420; Variance 107.281; scale 0.451

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description	Pred.	No.
1	1374	76.5	183	1	CORA_HPBVT	CORE ANTIGEN.	2.01e-22		
2	1358	75.6	211	1	CORA_HPBVA	CORE ANTIGEN.	2.15e-22		
3	1349	75.1	183	1	CORA_HPBVT	CORE ANTIGEN.	1.09e-22		
4	1336	74.4	183	1	CORA_HPBVO	CORE ANTIGEN.	3.14e-21		
5	1334	74.3	183	1	CORA_HPBVT	CORE ANTIGEN.	7.52e-21		
6	1332	73.6	183	1	CORA_HPBVT	CORE ANTIGEN.	1.40e-21		
7	1331	73.6	185	1	CORA_HPBVT	CORE ANTIGEN.	2.17e-21		
8	1314	73.2	195	1	CORA_HPBVT	CORE ANTIGEN.	4.57e-21		
9	1313	73.1	212	1	CORA_HPBVT	CORE ANTIGEN.	7.06e-21		
10	1308	72.8	183	1	CORA_HPBVT	CORE ANTIGEN.	6.23e-21		
11	1305	72.7	185	1	CORA_HPBVT	CORE ANTIGEN.	2.30e-21		
12	1304	72.6	214	1	CORA_HPBVT	CORE ANTIGEN.	3.56e-21		
13	938	52.2	188	1	CORA_HPBVT	CORE ANTIGEN.	2.07e-14		
14	934	52.0	217	1	CORA_HPBGS	CORE ANTIGEN.	1.56e-14		
15	933	51.9	187	1	CORA_HVB8	CORE ANTIGEN.	1.77e-14		
16	200	11.1	305	1	CORA_HPBHE	CORE ANTIGEN.	1.17e-13		
17	200	11.1	305	1	CORA_HPBDU	CORE ANTIGEN.	1.17e-13		
18	200	11.1	305	1	CORA_HPBWD	CORE ANTIGEN.	1.17e-13		
19	200	11.1	305	1	CORA_HPBDC	CORE ANTIGEN.	1.17e-13		
20	200	11.1	305	1	CORA_HPBDB	CORE ANTIGEN.	1.17e-13		
21	171	9.5	289	1	CERP_MSAU	CHOLESTERYL ESTER TRAN	2.27e-05		
22	171	9.5	493	1	CETP_MACFA	CHOLESTERYL ESTER TRAN	2.27e-05		
23	171	9.5	493	1	CETP_HOMAN	CHOLESTERYL ESTER TRAN	2.27e-05		

24	167	9.3	497	1	CETP_RABIT	CHOLESTERYL ESTER TRAN	8.49e-07
25	156	8.8	196	1	SPR2_CAEEL	PUTATIVE SPLICING FACT	1.59e-09
26	148	8.2	208	1	XSX2_CAEEL	HYPOTHETICAL 24.0 KD P	3.80e-06
27	136	7.6	78	1	PRTI_SEEOF	SERNATID-SPECIFIC PRO	1.52e-04
28	133	7.4	567	1	GPV_MOUSE	PLATELET GLYCOPROTEIN	3.75e-04
29	130	7.3	61	1	HSP1_MACEU	SPERM PROTAINE P1.	9.79e-04
30	130	7.2	77	1	PRT2_SEO	SERNATID-SPECIFIC PRO	9.12e-04
31	125	7.0	60	1	HSP1_MACGI	SPERM PROTAINE P1.	3.93e-03
32	125	7.0	61	1	HSP1_MACGI	SPERM PROTAINE P1.	3.93e-03
33	126	7.0	498	1	VE2_HPVO8	REGULATOR PROTEIN E2.	2.94e-03
34	126	7.0	1523	1	SON_HUMAN	SON PROTEIN (SON3).	2.94e-03
35	124	6.9	57	1	HSP1_IDDMA	SPERM PROTAINE P1.	5.24e-03
36	122	6.8	91	1	PHI_MYED	SEERM-SPECIFIC PROTEIN	9.29e-03
37	122	6.8	132	1	PR1_ANTER	SPERM-SPECIFIC PROTEIN	9.29e-03
38	121	6.7	61	1	HSP_CAEEL	SPERM HISTONE (PROTAM	1.23e-02
39	121	6.7	739	1	DDI3_CAEEL	PUTATIVE PRE-MNNA SPLI	1.23e-02
40	121	6.7	843	1	CYPI_BROWA	PETIDYLPYROLYL ISOMERA	1.23e-02
41	119	6.6	50	1	HSP1_RAT	SPERM PROTAINE P1 (CY	2.17e-02
42	119	6.6	50	1	HSP1_MOUSE	SPERM PROTAINE P1 (CY	2.17e-02
43	118	6.5	238	1	SFR7_HUMAN	SPLICING FACTOR, ARGIN	2.87e-02
44	118	6.5	576	1	CATA_RHOCA	PEOXIDASE / CATALASE	2.87e-02
45	115	6.4	102	1	HSP2_MACMU	SPERM HISTONE P2 PRECU	6.60e-02

ALIGNMENTS

ID	RESULT				
CORA_HPBVY	1	STANDARD;	PRT:	183 AA.	
AC	PO3146;				
DT	21-JUL-1986	(REL. 01,	CREATED)		
DT	21-JUL-1986	(REL. 01,	LAST SEQUENCE UPDATE)		
DT	01-APR-1990	(REL. 14,	LAST ANNOTATION UPDATE)		
DE	CORE ANTIGEN.				
GN	C.				
OS	HEPATITIS B VIRUS (SUBTYPE AYW).				
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE: 81012091.				
RA	GALIBERT F., MANDART E., FITOUSSI F., TIOLLAIS P., CHARNEY P.;				
RT	"Nucleotide sequence of the hepatitis B virus genome (subtype ayw)"				
RT	cloned in E. coli."				
RL	NATURE 281:646-650(1979).				
RN	[2]				
RP	SEQUENCE FROM N.A. (CLONE PHB320).				
RA	MEDLINE: 85204397.				
RA	BICHKO V., PUSHKO P., DREILINA D., PUMPEIN P., GREN E.;				
RT	"Subtype ayw variant of hepatitis B virus. DNA primary structure analysis"				
RT	FEBRS LETT. 185:208-212(1985).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpost - CC between the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licenses@isb-sib.ch).				
CC	-----				
DR	EMBL: X01460; G62278; ALT_INIT.				
DR	EMBL: X02496; -: NOT_ANNOTATED_CDS.				
DR	PIR: A03711; NKVLAH.				
DR	PIR: A03712; NKVLAH.				
DR	PRFM: PF00906; Hepatitis_core. 1.				
KM	CORE PROTEIN; REPEAT.				
FT	REPEAT 162 169				
FT	REPEAT 170 177				
FT	CONFICT 33 33 T -> N (IN REF. 2).				
FT	CONFLICT 80 80 A -> I (IN REF. 2).				
SQ	SEQUENCE 183 AA; 2116 MM; 2AE7A17A CRC32;				
Query Match	76.5%;	Score 1374;	DB 1;	Length 183;	
Best Local Similarity	100.0%;	Pred. No. 2,01e+226;			

Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAIL 60
 OY 58 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAIL 117
 Db 61 CWGELMTLATWGVNLEDPASRDLYSVYNTNMGKFKROLMLFHISCLTFGRETYEYLV 120
 OY 118 CWGELMTLATWGVNLEDPASRDLYSVYNTNMGKFKROLMLFHISCLTFGRETYEYLV 177
 Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRSOSRE 180
 OY 178 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRSOSRE 237
 Db 181 SOC 183
 OY 238 SOC 240

RESULT 2
 ID CORA_HPEVZ STANDARD: PRT: 211 AA.

AC P24023;
 DT 01-MAR-1992 (REL. 21, CREATED)
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
 DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.

OS HEPATITIS B VIRUS (STRAIN ALPHA1).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
 RX SEQUENCE FROM N.A.

RA MEDLINE: 90266476.
 RA TONG S., LI J., VITVITSKI L., TREPO C.;

RT Active hepatitis B virus replication in the presence of anti-Hbe is associated with viral variants containing an inactive pre-C region.";
 RL VIROLOGY 176:596-603(1990).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: M32138; G495034; ALT_SEQ.

DR PIR: A34773; NKVLAI.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT

FT REPEAT 183 190

FT REPEAT 198 206

SO SEQUENCE 211 AA; 24208 MW; 0924CE5E CRC32;

Query Match 75.6%; Score 1358; DB 1; Length 211;
 Best Local Similarity 95.2%; Pred. No. 2.15e-223;

Matches 179; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 24 LGWLMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTAL 83
 OY 53 LGSLMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTAL 112

Db 84 ROAILCWDGLITLSTWGVNLEDPTRSDLYSVYNTNMGKFKROLMLFHISCLTFGRETY 143
 OY 113 ROAILCWDGLITLSTWGVNLEDPTRSDLYSVYNTNMGKFKROLMLFHISCLTFGRETY 172

Db 144 IEVLVSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRR 203
 OY 173 IEVLVSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRR 232

Db 204 TOSRESOC 211
 OY 233 SOSRESOC 240

RESULT 3
 ID CORA_HPEVZ STANDARD: PRT: 183 AA.

AC P03147;
 DT 21-JUL-1986 (REL. 01, CREATED)

DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)

DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

GN C.

OS HEPATITIS B VIRUS (SUBTYPE ADYW).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
 RX SEQUENCE FROM N.A.

RA MEDLINE: 81012115.

RA PASER M., GOTO T., GILBERT W., ZINK B., SCHALLER H., MCKAY P.,
 RA LEADBETTER G., MURRAY K.;

RT "Hepatitis B virus genes and their expression in E. coli.";
 RL NATURE 282:575-579(1979).

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

CC EMBL: J02202; G329638; -

DR EMBL: A08967; G411874; -

DR PIR: B93217; NKVLAI.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT

FT REPEAT 162 169

FT REPEAT 170 177

SO SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;

Query Match 75.1%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 1.09e-221;

Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAIL 60
 OY 58 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAIL 117

Db 61 CWGELMTLATWGVNLEDPASRDLYSVYNTNMGKFKROLMLFHISCLTFGRETYEYLV 120
 OY 118 CWGELMTLATWGVNLEDPASRDLYSVYNTNMGKFKROLMLFHISCLTFGRETYEYLV 177

Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRSOSRE 180
 OY 178 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSOSRRRSOSRE 237

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240

Db 181 SOC 183
 OY 238 SOC 240


```

RL J. GEN. VIROL. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
CC -----
DR EMBL; D00330.; NOT_ANNOTATED_CDS.
DR PIR; B28925; NKVLJ2.
DR PFAM; PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21081 MW; 115BD9E3 CRC32;
Query Match 74.4%; Score 1336; DB 1; Length 183;
Best Local Similarity 96.2%; Pred. No. 3,14e-219;
Matches 176; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Db 1 MDIDYKEFGASVELSTFLPSDFPFVSVDLDLTASALYREALSEPHKSPHTALROAIL 60
Oy 58 MDIDYKEFGAVELLSFLPSDFPFVSVDLDLTASALYREALSEPHKSPHTALROAIL 117
Db 61 CMGELMNLATWGSNLEDPASRELVSYSVNMGLKIRQLMFHISCLTFGEIVLEYLY 120
Oy 118 CMGELMTLATWGVNLDEPDASHDLVVSYNTMGILKFOLLMFHSCLTFGEIVLEYLY 177
Db 121 SFGVMIRTPPAPRPNPAPILSTLPETTVARRGRSPPRRTPSPRRRSQSPRRRSQSRE 180
Oy 178 SFGVMIRTPPAPRPNPAPILSTLPETTVARRGRSPPRRTPSPRRRSQSPRRRSQSRE 237
Db 181 SOC 183
Oy 238 SOC 240
RESULT 5
ID CORA_HPBV4 STANDARD: PRT: 183 AA.
AC P03150; P03151;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DC CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADRI), HEPATITIS B VIRUS (SUBTYPE ADRA4),
OS AND HEPATITIS B VIRUS (SUBTYPE ADM / STRAIN INDOANSIA/PIM420) .
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ADR;
RX MEDLINE: 83168919.
RA ONO Y., OMDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RL DNA: subtype adr and adv." ;
RM NUCLEIC ACIDS RES. 11:1747-1757(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ADRA4;
RX MEDLINE: 83246570.
RA FUJIYAMA A., MIYANOHARA A., NOZAKI C., YONEYAMA T., OHOTO N.,
RA MATSUUBARA K.;
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RM adt." ;
RL NUCLEIC ACIDS RES. 11:4601-4610(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ADV;
RX MEDLINE: 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEMIGNJO R.I., IMAI M.,
RA MIYAKAWA Y., MAYUMI M.;
```

```

RT      "Typing hepatitis B virus by homology in nucleotide sequence:
RL      comparison of surface antigen subtypes.";
RR      J. GEN. VIROL. 69:2575-2583(1988).
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC
DR      EMBL; V00867; ? NOT_ANNOTATED_CDS.
DR      EMBL; X01587; G59407; -.
DR      EMBL; D00331; ? NOT_ANNOTATED_CDS.
DR      PIR; A83480; NKVLAS.
DR      PIR; B83450; NKVLAA.
DR      PIR; C28925; NKVLJ3.
DR      PFAM; PF00906; Hepatitis_core; 1.
KW      CORE PROTEIN; REPEAT.
FT      REPEAT              162   169
          REPEAT            170   177
SQ      SEQUENCE           183 AA; 21095 MW; D2E185B9 CRC32;
Query Match               74.3%; Score 1334; DB 1; Length 183;
Best Local Similarity     95.6%; Pred. No. 7,52e+219;
Matches    175; Conservative    5; Mismatches    3; Indels    0; Gaps    0;

Db         1 MDIDYKEFGAVELSTLPSPFFPSINDLDLTASALVREALSESRHSPHTALROAIL 60
          |||||
Qy         58 MDIDYKEFGAVELSTLPSPDFPSVDLDTASALVREALSESRHSPHTALROAIL 117
          |||||
Db         61 CGGEMLNATWGSNLEDPASRELIVSVNVNMGLKIKQLMFHSCLTFGRETVLEYLY 120
          |||||
Qy        118 CGGEMLNATWGVNLDEPDASDLVSVYNVMGAKFQLMLFHSLCTFGRETVIEIYL 177
          |||||
Db        121 SFGWIRTPPAPRPNPAPILSTLPETTVVRNRGRSPRRRTPSPRRRSQSPPRRRSQGRE 180
          |||||
Qy        178 SFGWIRTPPAPRPNPAPILSTLPETTVVRNRGRSPRRRTPSPRRRSQSPPRRRSQGRE 237
          |||||
Db         181 SOC 183
          |||
Qy        238 SOC 240

RESULT      6 STANDARD: PRT; 183 AA.
AC          p1791;
DT          01-AUG-1990 (REL. 15, CREATED)
DT          01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DI          01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE          CORE ANTIGEN*
OS          HEPATITIS B VIRUS (SUBTYPE ADM / STRAIN JAPAN/PUDM233).
OC          VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE: 89010694.
RA          OKAMOTO H., TSUDA F., SAKUGAWA H., SASIROSEMIGNIO R.I., IMAI M.,
RA          MIYAKAWA T., MATSUMI M.;
RT          "Typing hepatitis B virus by homology in nucleotide sequence:
RT          comparison of surface antigen subtypes.";
RL          J. GEN. VIROL. 69:2575-2583(1988).
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb.sib.ch).
CC
DR      EMBL; D00329; ? NOT_ANNOTATED_CDS.
DR      PIR; A28925; NKVLJ1.
```

DR PFAM: PF00906; Hepatitis-core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 21224 MW: F6B348B6 CRC32:

Query Match 73.6%; Score 1322; DB 1; Length 183;
Best local Similarity 96.2%; Pred. No. 1,40e-216;
Matches 176; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROAIL 60
|||||
QY 58 MDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROAIL 117
|||||
Db 61 CMGELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 120
|||||
QY 118 CMGELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 177
|||||
Db 121 SFQVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 180
|||||
QY 178 SFQVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 237
|||||
Db 181 SOC 183
|||||
QY 238 SOC 240

RESULT 7
ID CORA_HPBVT STANDARD; PRT: 185 AA.

AC P03149;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW)
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 83168919.
RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adw.";
RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: V00866; -! NOT_ANNOTATED_CDS.
DR PIR: C93460; NKVLA6.
DR PFAM: PF00906; Hepatitis-core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 164 171
FT REPEAT 172 179
SQ SEQUENCE 185 AA: 21394 MW: 791E0381 CRC32:

Query Match 73.6%; Score 1321; DB 1; Length 185;
Best local Similarity 96.2%; Pred. No. 2.17e-216;
Matches 178; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Db 1 MDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROAIL 60
|||||
QY 58 MDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROAIL 117
|||||
Db 61 CMGELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 120
|||||
QY 118 CMGELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 177
|||||

Db 121 SFQVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 180
|||||
QY 178 SFQVWIRTPPAYRPPNAPILSTLPETTVRRR--GSPRRRTPSPRRRSQSPRRRSQSR 235
|||||
Db 181 RESOC 185
|||||
QY 236 RESOC 240

RESULT 8
ID CORA_HPBVT STANDARD; PRT: 195 AA.

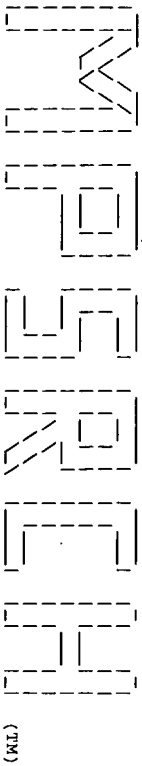
AC P29178;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF)
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 90169850.
RA BHAT R.A., ULRICH P.P., VYAS G.N.;
RT "Molecular characterization of a new variant of hepatitis B virus in
RT a persistently infected homosexual man.";
RL HEPATOLOGY 11:271-276(1990).
DR PIR: A37182; NKVLH3.
DR PFAM: PF00906; Hepatitis-core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 174 181
FT REPEAT 182 189
SQ SEQUENCE 195 AA: 22461 MW: AF3DB5F3 CRC32:

Query Match 73.2%; Score 1314; DB 1; Length 195;
Best local Similarity 93.5%; Pred. No. 4.57e-215;
Matches 173; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 11 FGIDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROA 70
|||||
QY 56 LSMDIDPKKEGATVLLSLPDSFPPSRDLDTASALYREALSPHCSPHHTALROA 115
|||||
Db 71 ILGWELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 130
|||||
QY 116 ILGWELMTLATVGNLDPASRDLYVNYVTNMGKIRQLMLFHSICLTFTGRETVEYL 175
|||||
Db 131 LVSEGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 190
|||||
QY 176 LVSEGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTPSPRRRSQSPRRRSQSR 235
|||||
Db 191 RESOC 195
|||||
QY 236 RESOC 240

RESULT 9
ID CORA_HPBVT STANDARD; PRT: 212 AA.

AC Q05495;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE W4B).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 93346970.
RA NAUMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,
RA GERLICH W.H.;
RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT that expresses HBV surface antigen subtype adw4.";
RL J. GEN. VIROL. 74:1627-1632(1993).
CC



Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:14:23 1999; Maspar time 17.75 Seconds
Tabular output not generated.

Title: >US3835-10-38183
Description: (1.240) from us3835-10-38183.pep
Perfect Score: 1796
Sequence: 1 MDIPYKEFGATVELLSFLP.....RRRSQSPRRRSQSPRESQC 240

Scoring table:
PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

sptremble9
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phase 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 46.641; Variance 107.390; scale 0.434

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	76.6	212 14	089656	PRE-C/CORE.	9.01e-219
2	1374	76.5	212 14	068020	PRE-C/CORE.	1.37e-218
3	1374	76.5	212 14	089597	HBCAG.	1.37e-218
4	1373	76.4	183 14	089437	X, PREC AND C GENES (C	2.09e-218
5	1373	76.4	183 14	068008	X, PREC AND C GENES (F	2.09e-218
6	1373	76.4	212 14	067876	PRE C/C ORF.	2.09e-218
7	1372	76.4	212 14	011884	CORE ANTIGEN PRECURSOR	3.18e-218
8	1370	76.3	183 14	068066	X, PREC AND C GENES (L	7.37e-218
9	1370	76.3	212 14	068068	PRE-C/CORE.	7.37e-218
10	1370	76.3	212 14	068025	PRE-C/CORE.	7.37e-218
11	1369	76.2	212 14	067872	PRE C/C ORF.	1.12e-217
12	1369	76.2	212 14	067984	PRE-C/CORE.	1.12e-217
13	1369	76.2	212 14	068032	PRE-C/CORE.	1.12e-217
14	1368	76.2	212 14	068010	PRE-C/CORE.	1.12e-217
15	1366	76.1	212 14	068070	PRE-C/CORE.	3.96e-217
16	1366	76.1	212 14	068030	PRE-C/CORE.	3.96e-217
17	1364	75.9	183 14	067989	X, PREC AND C GENES (M	9.17e-217
18	1363	75.9	183 14	068048	X, PREC AND C GENES (F	1.40e-216
19	1363	75.9	212 14	068014	PRE-C/CORE.	1.40e-216
20	1363	75.9	212 14	068012	PRE-C/CORE.	1.40e-216

21	1363	75.9	212 14	067980	PRE-C/CORE.	1.40e-216
22	1362	75.8	212 14	068051	PRE-C/CORE.	2.13e-216
23	1360	75.7	212 14	089446	PRE-C/CORE.	4.92e-216
24	1360	75.7	212 14	068016	PRE-C/CORE.	4.92e-216
25	1360	75.7	212 14	096845	COMPLETE GENOME.	4.92e-216
26	1359	75.7	212 14	068053	PRE-C/CORE.	7.50e-216
27	1357	75.6	212 14	068075	PRE-C/CORE.	1.74e-215
28	1357	75.6	212 14	068077	PRE-C/CORE.	1.74e-215
29	1355	75.4	212 14	068045	PRE-C/CORE.	4.02e-215
30	1352	75.3	183 14	089531	CORE PROTEIN.	1.42e-214
31	1351	75.2	183 14	067946	C ANTIGEN.	2.16e-214
32	1351	75.2	212 14	089719	PRE-C/CORE.	2.16e-214
33	1349	75.1	183 14	067997	X, PREC AND C GENES (B	5.01e-214
34	1348	75.1	183 14	068037	X, PREC AND C GENES (D	7.62e-214
35	1348	75.1	183 14	068030	X, PREC AND C GENES (M	7.62e-214
36	1347	75.0	212 14	067912	PRE-C/CORE PROTEIN.	1.16e-213
37	1345	74.9	183 14	067973	X, PREC AND C GENES (T	2.69e-213
38	1343	74.8	183 14	068023	X, PREC AND C GENES (C	6.12e-213
39	1344	74.8	212 14	092918	PRE-C/CORE.	4.09e-213
40	1344	74.8	212 14	068035	PRE-C/CORE.	4.09e-213
41	1342	74.7	183 14	067964	X, PREC AND C GENES (F	9.47e-213
42	1341	74.7	183 14	068064	X, PREC AND C GENES (S	1.44e-212
43	1341	74.7	183 14	068003	X, PREC AND C GENES (S	1.44e-212
44	1341	74.7	212 14	068042	PRE-C/CORE.	1.44e-212
45	1341	74.7	212 14	081115	HBEAG,HBCAG.	1.44e-212

ALIGNMENTS

RESULT 1
ID 089656; PRELIMINARY; PRT; 212 AA.
AC 089656;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHENI-1'85.
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RP SEQUENCE FROM N.A.
RC STRAIN-AYMA;
RA PLUCIENNICZAK A.;
RL SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 81012091.
RT GALBERT F., MANDART E., FINOUSSI F., TIOLLAIS P., CHARNAY P.;
RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayv)
RT cloned in E. coli.";
RL NATURE 281:646-650(1979).
[4]
RP SEQUENCE FROM N.A.
RA BORISOVA G.P., POMPEN P.P., BYCHKO V.V., PUSHKO P.M., KALIS Y.V.,
RA DISHLER A.V., GREN E.Y., TSIBINOGIN V.V., KUKAIN R.A.;
RL DOCL. BIOCHEM. 279:386-390(1985).
DR EMBL: X85290; G736116; -;
DR EMBL: X85300; G736152; -;
DR EMBL: X85313; G736196; -;
DR EMBL: X85283; G736090; -;
DR EMBL: X85306; G736174; -;
DR EMBL: 235716; G527437; -;
DR EMBL: J02203; G329644; -;
DR EMBL: X85312; G736193; -;
DR PFAM: PF00906; HepatitisB_core; 1.
KW SIGNAL.
SQ SEQUENCE 212 AA; 24350 MW; 71EA2C82 CRC32;

Query Match 76.6%; Score 1375; DB 14; Length 212;
Best Local Similarity 99.5%; Pred. No. 9, 01e-219;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
:|||||
QY 57 SMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 116
:|||||

Db 89 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 148
:|||||
QY 117 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 176
:|||||

Db 149 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 177 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
:|||||

Db 209 ESQC 212
:|||||
QY 237 ESQC 240

RESULT 2
ID Q68020 PRELIMINARY; PRT: 212 AA.
AC Q68020;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPAITIS B VIRUS.
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG-1'85;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85291; G736119; -;
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 76.5%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1, 37e-218;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
:|||||
QY 57 SMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 116
:|||||

Db 89 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 148
:|||||
QY 117 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 176
:|||||

Db 149 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 177 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
:|||||

Db 209 ESQC 212
:|||||
QY 237 ESQC 240

RESULT 3
ID Q89597 PRELIMINARY; PRT: 212 AA.
AC Q89597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HBCAG.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SUB-TYPE AYW;

RX MEDLINE: 94079539.
RA PREISLER-ADAMS S., SCHLAYER M.J., PETERS T., HETTLER F., GEROK W.,
RA RASENCK J.;
RT "Sequence analysis of hepatitis B virus DNA in immunologically
RT negative infection."
RL ARCH. VIROL. 133:385-396(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X72702; G288930; -;
DR EMBL: X80925; E198084; -;
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24336 MW; 575D20BF CRC32;

Query Match 76.5%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1, 37e-218;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
:|||||
QY 57 SMDIDPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 116
:|||||

Db 89 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 148
:|||||
QY 117 LCMGELMTLATVGVNLEDPAARDLVSVVNTNMGILKFKQLMFIHISCLTFGREVIEYL 176
:|||||

Db 149 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 177 VSFGWIRTTPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
:|||||

Db 209 ESQC 212
:|||||
QY 237 ESQC 240

RESULT 4
ID Q89437 PRELIMINARY; PRT: 183 AA.
AC Q89437;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (CASTA 2).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTA-2'87;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85289; G736113; -;
DR EMBL: X80925; E198085; -;
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 76.4%; Score 1373; DB 14; Length 183;

Best Local Similarity 99.5%; Pred. No. 2.09e-218;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAIL 60
|||
QY 58 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAIL 117
|||
DB 61 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 120
|||
QY 118 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 177
|||
DB 121 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
|||
QY 178 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 237
|||

DB 181 SOC 183
|||
QY 238 SOC 240

RESULT 5
ID 068008 PRELIMINARY; PRT; 183 AA.
AC 068008:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (FERRACUT 1).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUT-1'89;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85284; G736098; -;
PRFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 183 AA; 21102 MW; BBE9AFBB CRC32;

Query Match 76.4%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 2.09e-218;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAIL 60
|||
QY 58 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAIL 117
|||
DB 61 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 120
|||
QY 118 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 177
|||
DB 121 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
|||
QY 178 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 237
|||

DB 181 SOC 183
|||
QY 238 SOC 240

RESULT 6
ID 067876 PRELIMINARY; PRT; 212 AA.
AC 067876:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CI, HBV SUBTYPE AYW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65258; G59436; -;
PRFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;

Query Match 76.4%; Score 1373; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.09e-218;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAI 88
|||
QY 57 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAI 116
|||
DB 89 LCWELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 148
|||
QY 117 LCWELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 176
|||
DB 149 VAFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
|||
QY 177 VAFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
|||

DB 209 ESOC 212
|||
QY 237 ESOC 240

RESULT 7
ID 011884 PRELIMINARY; PRT; 212 AA.
AC 011884:
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN PRECURSOR.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA RAO B.S., CASEY J.L., RINAUDO J.S., KORBA B.E.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U95551; G2182120; -;
PRFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;

Query Match 76.4%; Score 1372; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3.18e-218;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAI 88
|||
QY 57 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHNTALRQAI 116
|||
DB 89 LCWELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 148
|||
QY 117 LCWELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLWPHISCLTFGRTVIEYL 176
|||
DB 149 VAFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
|||
QY 177 VAFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
|||

DB 209 ESOC 212
|||
QY 237 ESOC 240

RESULT 8
ID 068066 PRELIMINARY; PRT; 183 AA.
AC 068066:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (LICHERI 2).
GN CORE.
OS HEPATITIS B VIRUS.

```
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2'87;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 6F38A3B CRC32;

Query Match
Best Local Similarity 98.4%; Score 1370; DB 14; Length 183;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 60
58 MDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 117
Db 61 CWDGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 120
118 CWDGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 177
Db 121 SFGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
178 SFGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 237
Db 181 SOC 183
238 SOC 240

RESULT 9
ID 068068 PRELIMINARY; PRT; 212 AA.
AC 068068:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3'90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 88
57 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 116
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 148
117 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 176
Db 149 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
177 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
Db 209 ESOC 212
237 ESOC 240

RESULT 10
ID 068025 PRELIMINARY; PRT; 212 AA.
```

```
AC 068025:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG'3;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 88
57 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 116
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 148
117 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 176
Db 149 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
177 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
Db 209 ESOC 212
237 ESOC 240

RESULT 11
ID 067872 PRELIMINARY; PRT; 212 AA.
AC 067872:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
DE HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C, HBSAG POSITIV, HBV SUBTYPE ATW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24348 MW; A6BD9DA1 CRC32;

Query Match
Best Local Similarity 98.4%; Score 1369; DB 14; Length 212;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 88
57 GMDIDPKKEFGATVELLSFLPSDFEPSVRDLDTASALYREALSEPHSCSPHHTALROAI 116
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 148
117 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGRETIVIELY 176
Db 149 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
177 VSGVWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 236
Db 209 ESOC 212
```

```
QY 237 ESQC 240

RESULT 12
ID 067984 PRELIMINARY: PRT: 212 AA.
AC 067984:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-PATIENT VITTORINA'92;
RC LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RA SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85256; G736052;
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24322 MM; 83D9780B CRC32;

Query Match 76.2%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1,12e-217;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 88
QY :|||||
57 SMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 116

Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
117 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 176

Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 177 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236

Db 209 ESQC 212
QY 237 ESQC 240

RESULT 13
ID 068032 PRELIMINARY: PRT: 212 AA.
AC 068032:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-PATIENT CHIGLINE-2'86;
RC LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RA SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85296; G736139;
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24308 MM; 69D87B53 CRC32;

Query Match 76.2%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 1,12e-217;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 88
QY :|||||
57 SMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 116

Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
117 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 176

Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 177 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236

Db 209 ESQC 212
QY 237 ESQC 240
```

```
QY 117 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 176
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 177 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236

Db 209 ESQC 212
QY 237 ESQC 240

RESULT 14
ID 068010 PRELIMINARY: PRT: 212 AA.
AC 068010:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-PATIENT FERRACUTI-2'90;
RC LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RA SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100;
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24292 MM; 04A4D12D CRC32;

Query Match 76.2%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 1,71e-217;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 88
QY :|||||
57 SMDIPYKEFGATVELLSFSPDFFPSVBDLDTASATYREALSEPHCSPHHTALROAI 116

Db 89 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 148
QY :|||||
117 LCWGLMTLATWGVNLEDPASRDVSVYNTMGLKFRQLMFHISCLTFGRETVIEYL 176

Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 177 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 236

Db 209 ESQC 212
QY 237 ESQC 240

RESULT 15
ID 068070 PRELIMINARY: PRT: 212 AA.
AC 068070:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN-PATIENT GORDO'84;
RC LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RA SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210;
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24376 MM; 80F52D0F CRC32;

Query Match 76.1%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3,96e-217;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

Db 29 GMDIDPYKRGATVELLSFLPSDFPSPVROLDLTASALYREALSPHHTALROAI 88
QY 57 SMDIDPYKRGATVELLSFLPSDFPSPVROLDLTASALYREALSPHHTALROAI 116
Db 89 LCMGELMTLATWGVNLEDPAASRDVVSNTNMGKFRQLMFHISCLTFGRETVIEYL 148
QY 117 LCMGELMTLATWGVNLEDPAASRDVVSNTNMGKFRQLMFHISCLTFGRETVIEYL 176
Db 149 VSEGVWIRTPPAIRPPNAPILITLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 177 VSEGVWIRTPPAIRPPNAPILITLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 236
Db 209 ESOC 212
QY 237 ESOC 240

```

Search completed: Thu Dec 16 13:16:01 1999
 Job time : 98 secs.

 M050242
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_dp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:20:19 1999; MasPar time 12.84 Seconds

Tabular output not generated. 460.333 Million cell updates/sec

Title: >US3869-29-38183

Description: (1-278) from us3869-29-38183.pep

Perfect Score: 2088

Sequence: 1 MDIDPYKEFGATVELLSFLP.....RRRRSQSPRRRSQSRSSOC 278

Scoring table: PAM 150

Searched: 170751 segs, 21266608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.381; Variance 171.697; scale 0.194

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	65.9	194	32	W50242	Hepatitis B virus pre
2	1375	65.9	212	32	W50250	Hepatitis B virus pre
3	1375	65.9	346	5	R27473	S12/core protein.
4	1363	65.3	184	1	P80959	Hepatitis B virus sub
5	1359	65.1	193	32	W50241	Hepatitis B virus pre
6	1358	65.0	183	32	W50251	Hepatitis B virus p21
7	1353	64.8	183	20	W09044	Hepatitis B virus cor
8	1349	64.6	183	1	R05635	Hepatitis B antigen.
9	1349	64.6	183	5	P00041	Sequence of core anti
10	1344	64.4	397	20	W09048	Plasmodium falciparum
11	1338	64.1	184	5	P00004	Sequence of core anti
12	1336	64.0	196	8	R40806	Hepatitis B core / PV
13	1337	64.0	208	8	R40808	Hepatitis B core / PV
14	1334	63.9	183	4	R40311	Hepatitis B virus core
15	1334	63.9	183	8	R40805	Hepatitis B virus core
16	1334	63.9	183	13	R68868	Hepatitis B virus pol

Result	ID	Score	Query Match	Length	ID	Description	Pred. No.
17	1334	63.9	183	12	R62869	Hepatitis B virus core	2.55e-97
18	1333	63.8	198	8	R40807	Hepatitis B core / IL	3.08e-97
19	1328	63.6	293	10	R52286	Pres1 full length cor	7.98e-97
20	1325	63.5	183	13	R68866	Hepatitis B virus pol	1.41e-96
21	1325	63.5	214	1	P90702	Deduced amino acid se	1.41e-96
22	1325	63.5	289	20	W09049	Plasmodium falciparum	1.41e-96
23	1324	63.4	185	6	R30861	Hepatitis B core anti	1.71e-96
24	1324	63.4	185	10	R55284	Deduced sequence of f	1.71e-96
25	1324	63.4	185	6	R30780	Native HBcAg protein.	1.71e-96
26	1324	63.4	185	6	R30780	Native HBcAg protein.	1.71e-96
27	1321	63.3	185	26	P30061	Core antigen of HBV.	3.02e-96
28	1320	63.2	183	13	R68869	Hepatitis B virus pol	3.66e-96
29	1317	63.1	185	6	R30781	Altered HBcAg protein	6.47e-96
30	1317	63.1	185	6	R31026	Altered HBcAg protein	6.47e-96
31	1314	62.9	183	13	R68867	Hepatitis B virus pol	1.14e-95
32	1313	62.9	183	13	R68870	Hepatitis B virus pol	1.38e-95
33	1311	62.8	185	6	R30784	Altered HBcAg protein	2.03e-95
34	1311	62.8	185	6	R31029	Altered HBcAg protein	2.03e-95
35	1311	62.8	185	6	R30865	Modified hepatitis B	2.03e-95
36	1309	62.7	185	6	R30862	Hepatitis B core anti	2.96e-95
37	1308	62.6	185	6	R30783	Altered HBcAg protein	3.58e-95
38	1308	62.6	185	6	R30864	Modified hepatitis B	3.58e-95
39	1302	62.4	203	1	P82872	HBcAg/beta-Gal fusion	1.12e-94
40	1302	62.4	214	1	P80961	HBV core antigen enco	9.28e-93
41	1287	61.6	198	32	W50252	Hepatitis B virus pre	1.95e-93
42	1287	61.6	198	32	W50252	Hepatitis B virus pre	1.87e-91
43	1263	60.5	183	17	R98878	Hepatitis B virus E a	3.99e-91
44	1259	60.3	183	17	R98883	Hepatitis B virus E a	4.88e-91
45	1258	60.2	183	17	R98884	Hepatitis B virus E a	4.88e-91

ALIGNMENTS

RESULT 1
 ID W50242 standard; Protein; 194 AA.

AC W50242:29-1998 (first entry)

DE Hepatitis B virus precore p22 polypeptide Met-p22.

KW Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;

KW hepatocyte; liver; Met-p22.

OS Hepatitis B virus.

OS Synthetic.

FT Key

FT Protein

FT Location/Qualifiers

FT 2..194

FT /label= p22

PD W09809649-A1.

PD 12-MAR-1998.

PF 03-SEP-1997; U15500.

PF 03-SEP-1996; US-025370.

PA (GENE) GEN HOSPITAL CORP.

PI Melegari M, Scaglioni PP, Wands JR;

DR WPI: 98-193325/17.

PT DNA encoding proteins which can be incorporated with wild type

PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for

PT inhibition of viral replication, especially hepatitis B virus

PS Claim 11; Page 40; 60pp; English.

CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)

CC protein with an added N-terminal Met residue. p22 is produced by

CC elimination of the 19-amino acid leader peptide from the 25 kDa

CC full-length HBV precore protein (see W50250). Evidence is provided

CC that HBV replication is inhibited in the presence of high levels of

CC HBV precore or precore-related proteins. These proteins can be

CC incorporated into HBV nucleocapsids along with the p21 core protein

CC (see W50251), which is the usual nucleocapsid component, and

CC thereby render the nucleocapsids deficient in encapsidating HBV

CC pregenomic RNA. Thus, over-expression of the precore proteins, or

CC certain variants of them, leads to transdominant inhibition of HBV

CC replication. Suitable inhibitory proteins include p25 (see W50250),

CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het

CC (see W50238). Heterologous peptides (see W50244-49) may be

CC inserted into the p22 and Met-p22 polypeptides. The inhibitory

CC proteins can be produced by recombinant methods using claimed

CC expression vectors and host cells. They can be provided exogenously

CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 194 AA:

Query Match 65.9%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 1.04e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmidpkykefgatvelisfipsgdfsvrldltasalyrealsepshcsphtalrqa1 70
 :|||||
 QY 95 SMDIDPKKEGATVELLSFIPSDFFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 154
 Db 71 lwcgelmrlatwgvnljedpasrdlvsyvtntmgjkfgrllwfhisccltfgretvleyl 130
 :|||||
 QY 155 LCGELMTLATWGVNLJEDPASRDVLVSYVTNMGLKFRQLMFMHISCLEFGRETVLEYL 214
 Db 131 vsfgvwlrtppayrppnaillstlpetvtvrrrgsrprrrtsqrrrrtsqrr 190
 :|||||
 QY 215 VSGVWIRTPPAYRPPNADILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
 Db 191 esgc 194
 :|||||
 QY 275 ESQC 278

RESULT 2
 ID W50250 standard; Protein: 212 AA.
 AC W50250:
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p25 polypeptide.
 KW Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
 KW hepatocyte; liver; p25 protein.
 OS Hepatitis B virus.
 PN MO9809649-A1.
 PD 12-MAR-1998.
 PE 03-SEP-1997; 015500.
 PR 03-SEP-1996; 05-025370.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR:
 DR WPI: 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 15: Page 35; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsitating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SQ Sequence 212 AA:

Query Match 65.9%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 1.04e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmidpkykefgatvelisfipsgdfsvrldltasalyrealsepshcsphtalrqa1 88
 :|||||
 QY 95 SMDIDPKKEGATVELLSFIPSDFFPSVRDLDTASALYREALSEPHCSPHHTALRQAI 154

Db 89 lwcgelmrlatwgvnljedpasrdlvsyvtntmgjkfgrllwfhisccltfgretvleyl 148
 :|||||
 QY 155 LCGELMTLATWGVNLJEDPASRDVLVSYVTNMGLKFRQLMFMHISCLEFGRETVLEYL 214

Db 149 vsfgvwlrtppayrppnaillstlpetvtvrrrgsrprrrtsqrrrrtsqrr 208
 :|||||
 QY 215 VSGVWIRTPPAYRPPNADILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 274

Db 209 esgc 212
 :|||||
 QY 275 ESQC 278

RESULT 3
 ID R27473 standard; Protein: 346 AA.
 AC R27473:
 DT 24-FEB-1993 (first entry)
 DE S12/core protein.
 KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KW vaccinia virus; I3L; promoter; NVVAC; recombinant; HBV L;
 KW large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
 KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KW deletion loci; recipient loci.
 OS Synthetic.
 FH Key
 FT region
 FT location/Qualifiers
 FT 1..108
 FT /label S1
 FT region 109..163
 FT /label S2
 FT 164..346
 FT /label Core
 PN WO9215672-A.
 PD 17-SEP-1992.
 PE 09-MAR-1992; 001906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRO-) VIROGENETICS CORP.
 PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Limbach KJ, Norton EK, Paoletti E, Perkins ME, Pincus SE,
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI: 92-331718/40.
 DR N-PSDB: Q29105.

PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure: Fig 13; 45pp; English.
 CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the 13L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NVVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, lpsAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NVVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NVVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC Q35501-864.
 SQ Sequence 346 AA:

Query Match 65.9%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 1.04e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 mmdidpkykefgatvelisfipsgdfsvrldltasalyrealsepshcsphtalrqa1 222

```

OY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 154
    :|||||
DB 223 lcgwelmrlatwgvnledpasrdlvsyvtntmgjkfrqlwfmiscltfgretvleyl 282
    :|||||
OY 155 LCGWELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFMHISCLTGTGRETIVYL 214
    :|||||
DB 283 vsfgvwlrtppayrppnapilsltpetlvrrrgsrprrrrrsgspprrrrsgsr 342
    :|||||
OY 215 VSFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRRSOSPRRRRSOSR 274
    :|||||
DB 343 esgc 346
    :|||||
OY 275 ESOC 278

RESULT 4
ID P80959 standard; protein: 184 AA.
AC P80959;
DT 19-NOV-1990 (first entry)
DE Hepatitis B virus subtype ayw. core protein.
KM Conjugate: fusion protein; hepatitis B core protein; HBV subtype ayw.
KW T cell stimulating polypeptide; vaccines.
OS Synthetic.
PN EP-271302-A.
PD 15-JUN-1988.
PF 07-DEC-1987; 310725.
PR 07-OCT-1987; US-106538.
PS 07-OCT-1987; US-939617.
PA (SCRT-) Scripps Clinic Res.
PI Thornton GB, Mortalcy AM, Millich DR, McLachlan A.;
DR WIR: 88-163287/24.
PT New conjugates and fusion proteins of immunogenic polypeptide -
PT and hepatitis B core antigen and T cell stimulating polypeptide
PT corep. to core antigen, useful in vaccines.
PS Disclosure: P: English.
CC This sequence contains the T cell stimulating epitopes, amino acid
CC residue 1-55 and 70-140. It is believed that the regions 1-44 and
CC 70-140 do not contain determinants that suppress T cell activation.
CC Polypeptides essentially consisting of 15 to 55 amino acids
CC corresponding to the above mentioned HBV regions are T cell
CC stimulating. Coupling a polypeptide immunogen to such sequences,
CC e.g. by using a bifunctional reagent which forms a disulphide link,
CC improves its immunogenicity. These are useful in vaccines and can be used
CC therapeutically to improve T cell response to HBcAg in infected
CC subjects.
CC See also P80896-P80898 and P80951-P80859.
SQ Sequence 184 AA;

Query Match 65.3%; Score 1363; DB 1; Length 184;
Best Local Similarity 99.5%; Pred. No. 1.02e-99;
Matches 183; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

DT 28-SEP-1998 (first entry)
DE Hepatitis B virus precore p22 polypeptide.
KM Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
KW hepatocyte; liver; p22 protein.
OS Hepatitis B virus.
FH Key
FT Region
    82..98 /note= "immunodominant region"
FT WO9809649-A1.
PD 12-MAR-1998.
PF 03-SEP-1997; U15500.
PR 03-SEP-1996; US-025370.
PI (GEHO) GEN HOSPITAL CORP.
PI Melegari M, Scaglioni PP, Wands JR;
DR WPI: 98-193325/17.
PT DNA encoding proteins which can be incorporated with wild type
PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
PT inhibition of viral replication, especially hepatitis B virus
PS Claim 9; Page 34-35; 60pp; English.
CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
CC protein that is produced by elimination of the 19-amino acid leader
CC peptide from the 28 kDa full-length HBV precore protein (see
CC W50250). Evidence is provided that HBV replication is inhibited in
CC the presence of high levels of HBV precore or precore-related
CC proteins. These proteins can be incorporated into HBV nucleocapsids
CC along with the p21 core protein (see W50251), which is the usual
CC nucleocapsid component, and thereby render the nucleocapsids
CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
CC expression of the precore proteins, or certain variants of them,
CC leads to transdominant inhibition of HBV replication. Suitable
CC inhibitory proteins include p25 (see W50250), p22, Met-p22 (see
CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
CC (see W50238). Heterologous peptides (see W50244-49) may be
CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
CC proteins can be produced by recombinant methods using claimed
CC expression vectors and host cells. They can be provided exogenously
CC to the target cells for use in inhibiting HBV replication.
CC Alternatively, a nucleic acid construct that directs overexpression
CC of an inhibitory protein in target cells is used for the gene
CC therapy of HBV infection.
SQ Sequence 193 AA;

Query Match 65.1%; Score 1359; DB 32; Length 193;
Best Local Similarity 98.4%; Pred. No. 2.18e-99;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```


PF 1-JAN-1989; 123526.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOJ) Biogen Inc.
 PI Murray K, Schaller HE;
 DR WPI: 90-195067/26.
 DR N-PSDB: 004799
 PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
 PS disclosed in detection of infection and in vaccine prodn.
 CC Polypeptide fragments encoded by the DNA sequence are antigenic for
 CC HBV and may be used as a vaccine or in detection. Peptides may be
 CC cultured in a suitable bacterial host such as E.coli.
 CC Fragments of the sequence are also claimed as being antigenically
 CC useful.
 SO Sequence 183 AA;

Query Match 64.6%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 1,476-98;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdipkyefgatlvelsfpsdfpsvrdldtaaalrydalespeshphltalrgail 60
 |||
 QY 96 MDIPYKFGATVELSLPSDFPSVVDLDTASALYREALSEPHESPHHTALRQAIL 155
 |||
 Db 61 cwgdlmtlatwgvnlgedpaardlvsvyvtngvjkfqqllwfhscitfgretvleylv 120
 |||
 QY 156 CWGELMTLATWGVNLGLEDPASRDLVSVYNTNMGKLFQRLMFHISCLTFGRETVIEYLV 215
 |||
 Db 121 sfgywlrtpayrpnpnallstlpctvrrrgsprrrrrpsrrrrsgsrr 180
 |||
 QY 216 SFGWIRTPAYRPNAPILSTLPETVVRKRGSPRRRTSPRRRSQSPRRRSQSRE 275
 |||
 Db 181 sqc 183
 |||
 QY 276 SQC 278

RESULT 9
 ID P00041 standard; Protein; 183 AA.
 AC P00041;
 DT 14-OCT-1992 (first entry)
 DE Sequence of core antigen.
 KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
 OS Hepatitis B virus.
 PN EP-13828-A.

PD 06-AUG-1980.
 PF 21-DEC-1979; 303017.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOJ) Biogen Inc.
 PI Murray K, Schaller HE;
 DR WPI: 80-57268C/33.
 DR N-PSDB: N000003.
 PT Recombinant DNA coding for polypeptide - have specificity of
 PT hepatitis B viral antigens in detection or antibody stimulation
 PS Example; Figs 3-4; 43pp; English.
 CC Human serum from a single HBsAg positive, HBeAg positive donor
 CC (serotype adpm) was used to prep. a DNA-contg. pellet which was
 CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
 CC The labelled DNA was then extracted with phenol from the resulting
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
 CC It was then cloned in plasmid pBR322 which was used to transform E.
 CC coli. Micro-organisms prep. by the processes are deposited at the
 CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dG:
 CC HBV-Xba I dG; Tetr Amps HBV+.
 SO Sequence 183 AA;

Query Match 64.6%; Score 1349; DB 5; Length 183;
 Best Local Similarity 96.7%; Pred. No. 1,476-98;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdipkyefgatlvelsfpsdfpsvrdldtaaalrydalespeshphltalrgail 60
 |||
 QY 96 MDIPYKFGATVELSLPSDFPSVVDLDTASALYREALSEPHESPHHTALRQAIL 155
 |||
 Db 61 cwgdlmtlatwgvnlgedpaardlvsvyvtngvjkfqqllwfhscitfgretvleylv 120
 |||
 QY 156 CWGELMTLATWGVNLGLEDPASRDLVSVYNTNMGKLFQRLMFHISCLTFGRETVIEYLV 215
 |||
 Db 121 sfgywlrtpayrpnpnallstlpctvrrrgsprrrrrpsrrrrsgsrr 180
 |||
 QY 216 SFGWIRTPAYRPNAPILSTLPETVVRKRGSPRRRTSPRRRSQSPRRRSQSRE 275
 |||
 Db 181 sqc 183
 |||
 QY 276 SQC 278

RESULT 10
 ID W09048 standard; Protein; 397 AA.
 AC W09048;
 DT 11-APR-1997 (first entry)
 DE Plasmid pHBV DN AA encoded HBV dominant negative polypeptide.
 KW Hepadnavirus; woodchuck hepatitis virus; hepatitis delta virus;
 KW HBV; ground squirrel hepatitis B virus; duck hepatitis B virus;
 KW core protein; replication; antiviral; gene therapy; pHBV DN AA.
 OS Hepatitis B virus.
 FH Key
 FT Location/Qualifiers
 FT 1..179
 FT /label=HBV_core
 FT /note="positions 1-179 correspond to amino acids
 FT 1-179 of HBV core protein."
 FT 180..397
 FT /note="positions 180-397 correspond to amino
 FT acids 9-226 of HBV surface protein"

FT region
 FT W09700698-A1.
 PD 09-JAN-1997.
 PF 20-JUN-1996; U10602.
 PR 20-JUN-1995; US-017814.
 PA (GENO.) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 97-08716/08.
 DR N-PSDB: T49598.
 PT New method for inhibiting the replication of hepadnaviruses -
 PT comprises introducing a mutant polypeptide with a mutated core
 PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
 PT B

PS Disclosure; Page 46-48; 83pp; English.
 CC Plasmid pHBV DN AA carries an insert (T49598) coding for a protein
 CC (W09048) comprising a hepatitis B virus (HBV) core protein (see
 CC also W09044) fused in-frame at amino acid 179 with the HBV surface
 CC protein (see also W09045) at amino acid 8. Plasmid pHBV DN BB
 CC (T49599) expresses the HBV core fused at amino acid 175 to the
 CC surface protein at amino acid 112 (W09049). pHBV DN AA was at
 CC least as potent an inhibitor of HBV replication as construct
 CC pHBV DN (T49597); pHBV DN BB was less inhibitory than pHBV DN.
 CC Vectors expressing hepadnavirus dominant negative core mutants can
 CC be utilized in the gene therapy of viral infections.
 SO Sequence 397 AA;

Query Match 64.4%; Score 1344; DB 20; Length 397;
 Best Local Similarity 100.0%; Pred. No. 3,808-98;
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdipkyefgatlvelsfpsdfpsvrdldtaaalrydalespeshphltalrgail 60
 |||
 QY 96 MDIPYKFGATVELSLPSDFPSVVDLDTASALYREALSEPHESPHHTALRQAIL 155
 |||
 Db 61 cwgdlmtlatwgvnlgedpaardlvsvyvtngvjkfqqllwfhscitfgretvleylv 120
 |||
 QY 156 CWGELMTLATWGVNLGLEDPASRDLVSVYNTNMGKLFQRLMFHISCLTFGRETVIEYLV 215
 |||
 Db 121 sfgywlrtpayrpnpnallstlpctvrrrgsprrrrrpsrrrrsgsrr 179

QY 216 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274

RESULT 11

ID P00004 standard: Protein; 184 AA.
AC P00004;
DT 14-OCT-1992 (first entry)
DE Sequence of core antigen.
KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
OS Hepatitis B virus.
PN EP-13828-A.
PF 06-AUG-1980.
PR 21-DEC-1979; 303017.
PR 22-DEC-1978; GB-049907.
PR 27-DEC-1978; GB-050039.
PR 01-NOV-1979; GB-037910.
PA (BIOJ) BIOGEN NV.
PI Murray K, Schaller HE;
DR WPI; 80-57268C/33.
DR N-PSDB; N00002.
PT Recombinant DNA coding for polypeptide - have specificity of
PT hepatitis B viral antigens in detection or antibody stimulation
PS Claim 13; Page 40; 43pp; English.

CC Human serum from a single HBsAg positive, HBeAg positive donor
CC (serotype adym) was used to prep. a DNA-contg. pellet which was
CC labeled with 3H or 32P as described by P. M. Kaplan et al (1973).
CC The labeled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prepd. by the processes are deposited at the
CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst 1 dg;
CC HBV-Kpn I dc; Tetr Amps HBV+.
SQ Sequence 184 AA:

Query Match 64.1%; Score 1338; DB 5; Length 184;

Best Local Similarity 96.2%; Pred. No. 1.19e-97;
Matches 177; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Db 1 mdldpykefagavellsfpsdfpsvcrdlldtaaalrydalespeshphthalrgai 60
QY 96 MDIDPYKEFGATVELLSFSPDFPSVCRDLDTAALRYDALESPEHSPHHTALROAI 154
Db 61 lcvgdlnltatwgtntedpasrdlvsvyntnmgikfrqlwfhiscitfgretvleyl 120
QY 155 LCWGEIMLTATWGVNLEDPASRDLVSVYNTNMGKFRQLWFHISCITFGRETVIEYL 214
Db 121 vsfgvwtirtppayrppnapilstlpettvrrgrspprrrtsprrrsqsprrrrsqsr 180
QY 215 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
Db 181 esgc 184
QY 275 ESQC 278

RESULT 12
ID R40806 standard: Protein; 196 AA.
AC R40806;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 fusion.
KW Hepatitis B core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.
FH Key
FT region Location/Qualifiers
FT 7..13 /label= PV-1
FT 14..196 /label= HBC
FT protein
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
RA (NISW) NISSHIN OIL MILLS LTD.

DR WPI; 93-277479/35.

DR N-PSDB; Q47736.
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure: Fig 7; 12pp; Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
SQ Specification.
SQ Sequence 196 AA:

Query Match 64.0%; Score 1336; DB 8; Length 196;

Best Local Similarity 95.1%; Pred. No. 1.74e-97;
Matches 175; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Db 13 amldipykefagavellsfpsdfpsvcrdlldtaaalrydalespeshphthalrgai 72
QY 95 SMIDPYKEFGATVELLSFSPDFPSVCRDLDTAALRYDALESPEHSPHHTALROAI 154
Db 73 lcvgdlnltatwgtntedpasrdlvsvyntnmgikfrqlwfhiscitfgretvleyl 132
QY 155 LCWGEIMLTATWGVNLEDPASRDLVSVYNTNMGKFRQLWFHISCITFGRETVIEYL 214
Db 133 vsfgvwtirtppayrppnapilstlpettvrrgrspprrrtsprrrsqsprrrrsqsr 192
QY 215 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
Db 193 esgc 196
QY 275 ESQC 278

RESULT 13

ID R40806 standard: Protein; 208 AA.

AC R40806;
DT 16-FEB-1994 (first entry)

DE Hepatitis B core / PV-1 / IL-1 fusion.
KW Hepatitis B core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FH Key
FT region Location/Qualifiers
FT 8..19 /label= PV-1
FT 20..25 /label= IL-1
FT 26..208 /label= HBC

FT protein
FT J05192170-A.
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISW) NISSHIN OIL MILLS LTD.
DR WPI; 93-277479/35.
DR N-PSDB; Q47738.

PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure: Fig 9; 12pp; Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
SQ Specification.
SQ Sequence 208 AA:

Query Match 64.0%; Score 1337; DB 8; Length 208;

Best Local Similarity 94.6%; Pred. No. 1.44e-97;
Matches 175; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Db 24 lgmldipykefagavellsfpsdfpsvcrdlldtaaalrydalespeshphthalrgai 83
QY 94 LSMIDPYKEFGATVELLSFSPDFPSVCRDLDTAALRYDALESPEHSPHHTALROAI 153
Db 84 lcvgdlnltatwgtntedpasrdlvsvyntnmgikfrqlwfhiscitfgretvleyl 143

```

OY 154 ILGELMTLATWGVNLEDPASRDLVSYVNTNKGILFROLWPHISCLTGFRETIVLEY 213
DB 144 IYSGVWIRTPPAYRPPNAPILSTLPETVVRIGRSPTTSPRRIRSGSPRRIRSGS 203
OY 214 LVSGVWIRTPPAYRPPNAPILSTLPETVVRIGRSPTTSPRRIRSGSPRRIRSGS 273
DB 204 resgc 208
OY 274 RESQC 278

```

```

RESULT 14
ID P40311 standard; Protein; 183 AA.
AC P40311;
DT 18-AUG-1992 (first entry)
DE Hepatitis virus core antigen.
KM HBCAg; vaccine; diagnosis; HBV infection.
OS Hepatitis B virus.
PN J59074985-A.
PD 27-APR-1984.
PF 19-OCT-1982; 183432.
PR 19-OCT-1982; JP-183432.
PA (TAKE ) TAKEDA CHEMICAL IND KK.
DR WPI; 84-14321/23.
PT DNA used in prevention of infections by hepatitis virus B -
  comprises structural gene of hepatitis virus adr B surface antigen
  coding gene and at least 1 virus core antigen structural gene.
PS Disclosure; Fig 3; 13pp; Japanese.
CC The sequence is that of hepatitis virus core antigen (HBCAg). It
  can be used as a vaccine for the prevention of infections by
  hepatitis B virus (HBV) and also in the diagnosis of early stages
  of HBV infection. See also P40310.
SQ Sequence 183 AA.

```

```

Query Match 63.9%; Score 1334; DB 4; Length 183;
Best Local Similarity 95.6%; Pred. No. 2.55e-97;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdidpykefgasvvelsfpsdfffpsirdltdtasalyrealsephepcphhtalrqail 60
OY 96 MDIDPYKEFGATVELSLFSPDFPSVVDLDTASALYREALSEPHEPCSPHHTALRQAIL 155
DB 61 CWGELMTLATWGVNLEDPASRDLVSYVNTNKGILFROLWPHISCLTGFRETIVLEYLV 120
OY 156 CWGELMTLATWGVNLEDPASRDLVSYVNTNKGILFROLWPHISCLTGFRETIVLEYLV 215
DB 121 sfgywlrtppayrppnapilstlpetvvrigrspttsprrirsgspttsprrirsgsre 180
OY 216 SFGVWIRTPPAYRPPNAPILSTLPETVVRIGRSPTTSPRRIRSGSPRRIRSGSRE 275
DB 181 sqc 183
OY 276 SQC 278

```

```

RESULT 15
ID R40805 standard; Protein; 183 AA.
AC R40805;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core protein.
KM Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Hepatitis B virus.
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISW ) NISSHIN OIL MILLS LTD.
DR WPI; 93-277479/35.
N-PSDB: 047735.
PT Recombinant plasmid for high immunogenity virus - contains
  recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
  virus and exotic genes

```

```

PS Disclosure; Fig 6; 12pp; Japanese.
CC The hepatitis B core gene is recombined with PV-1 DNA
  and IL-1 beta to form a plasmid (047738) which is then used to produce
  a vaccine. NB: Sequence is difficult to read in the original
  CC specification.
SQ Sequence 183 AA;

```

```

Query Match 63.9%; Score 1334; DB 8; Length 183;
Best Local Similarity 95.6%; Pred. No. 2.55e-97;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdidpykefgasvvelsfpsdfffpsirdltdtasalyrealsephepcphhtalrqail 60
OY 96 MDIDPYKEFGATVELSLFSPDFPSVVDLDTASALYREALSEPHEPCSPHHTALRQAIL 155
DB 61 CWGELMTLATWGVNLEDPASRDLVSYVNTNKGILFROLWPHISCLTGFRETIVLEYLV 120
OY 156 CWGELMTLATWGVNLEDPASRDLVSYVNTNKGILFROLWPHISCLTGFRETIVLEYLV 215
DB 121 sfgywlrtppayrppnapilstlpetvvrigrspttsprrirsgspttsprrirsgsre 180
OY 216 SFGVWIRTPPAYRPPNAPILSTLPETVVRIGRSPTTSPRRIRSGSPRRIRSGSRE 275
DB 181 sqc 183
OY 276 SQC 278

```

```

Search completed: Thu Dec 16 13:23:42 1999
Job time : 203 secs.

```


US3869-29-38183
MDIDPYKEGATVELLSFLPSDFPVSVDLDTASALYREALSPHHTALROAILCWEGLMTL
DGLLLLOMDGFPFELHLYDFLOSISMDIDPYKEGATVELLSFLPSDFPVSVDLDTASALYREALSP
EHCSPHHTALROAILCWEGLMTLWGVNLEDPASRDVVSIVNTNGLKFRQLWPHISCLTFGRETV
IEXLVSGWIRTPPAYRPNNAPILSTLPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSRESOCI


```
##experimental_source isolate patient Licheri-1'85
#accession S53249
##molecule_type DNA
##residues 1-212 ##label LA5
##cross-references EMBL:X85306; NID:g736172; PID:g736174
##experimental_source isolate patient Flore-1'86
#accession S53262
##molecule_type DNA
##residues 1-212 ##label LA6
##cross-references EMBL:X85312; NID:g736191; PID:g736193
##experimental_source isolate patient Licheri'83
#accession S53277
##molecule_type DNA
##residues 30-212 ##label LA7
##cross-references EMBL:X85317; NID:g736211; PID:g736214
##experimental_source isolate patient Giorio-2'86
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be processed

REFERENCE A93214
#authors Galibert, F.; Mandart, E.; Fickoussi, F.; Tiollais, P.;
#journal Nature (1979) 281:646-650
#title Nucleotide sequence of the hepatitis B virus genome (subtype
#cross-references MUID:81012091
#accession A03711
##molecule_type DNA
##residues 1-212 ##label GAL
##cross-references GB:J02203; NID:g329640; PID:g329642
##experimental_source subtype ayw

GENETICS
#gene C
#classification #superfamily hepatitis B virus core antigen
#keywords alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link ENG #status
predicted #label ECP
SUMMARY #length 212 #molecular_weight 24350 #checksum 782

Query Match 65.9%; Score 1375; DB 1; Length 212;
Best Local Similarity 99.5%; Pred. No. 4.25e-205;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
:|||||
QY 95 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 154
:|||||
Db 89 LCMGELMTLATWGVNLEDPASRDLVVSYVNTNMGLEFRQLMFMHISCLTFGRTVIEYL 148
|||||
QY 155 LCMGELMTLATWGVNLEDPASRDLVVSYVNTNMGLEFRQLMFMHISCLTFGRTVIEYL 214
|||||
Db 149 VSEGVWIRTPPAYRPPNAPILSTPETTVARRGRSPRRRSPRRRSOSPRRRRSOSR 208
|||||
QY 215 VSEGVWIRTPPAYRPPNAPILSTPETTVARRGRSPRRRSPRRRSOSPRRRRSOSR 274
|||||
Db 209 ESOC 212
|||
QY 275 ESOC 278

RESULT 2
ENTRY S53207 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castaa-2'87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997
```

```
ACCESSIONS S53207
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53207
##molecule_type DNA
##residues 1-183 ##label LA1
##cross-references EMBL:X85289; NID:g736110; PID:g736113
##experimental_source isolate patient Castaa-2'87
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
#classification #superfamily hepatitis B virus core antigen
#keywords core protein
SUMMARY #length 183 #molecular_weight 21102 #checksum 2165

Query Match 65.8%; Score 1373; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 9.36e-205;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 60
:|||||
QY 96 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 155
:|||||
Db 61 CMEGELMTLATWGVNLEDPASRDLVVSYVNTNMGLEFRQLMFMHISCLTFGRTVIEYL 120
|||||
QY 156 CMEGELMTLATWGVNLEDPASRDLVVSYVNTNMGLEFRQLMFMHISCLTFGRTVIEYL 215
|||||
Db 121 SFGVWIRTPPAYRPPNAPILSTPETTVARRGRSPRRRSPRRRSOSPRRRRSOSRE 180
|||||
QY 216 SFGVWIRTPPAYRPPNAPILSTPETTVARRGRSPRRRSPRRRSOSPRRRRSOSRE 275
|||||
Db 181 SOC 183
|||
QY 276 SOC 278

RESULT 3
ENTRY S53211 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Castag-1'85 and others)
ALTERNATE_NAMES HBC antigen; Hbc antigen precursor / Hbc antigen; pre-C/C
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castag-1'85; isolate patient Ferracuti-1'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997
ACCESSIONS S53211; S53197
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53211
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85291; NID:g736117; PID:g736119
##experimental_source isolate patient Castag-1'85
#accession S53197
##molecule_type DNA
##residues 30-212 ##label LAW
##cross-references EMBL:X85284; NID:g736095; PID:g736098
##experimental_source isolate patient Ferracuti-1'89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
#classification #superfamily hepatitis B virus core antigen
#keywords alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label ENG\
```



```

08-Sep-1997
ACCESSIONS      S53270
REFERENCE        S53112
#authors         Lai, M.E.; Mazzeoli, A.P.; Porru, A.; Balestrieri, A.
#submission      submitted to the EMBL Data Library, March 1995
#accession       S53270
##molecule_type DNA
##residues       1-183 ##label LAI
##cross-references EMBL:X85314; NID:g736201; PID:g736204
##experimental_source isolate patient Licheri-2/87
#note            due to a stop codon between the alternative initiators
                  the e antigen precursor cannot be produced

GENETICS
#gene            C
CLASSIFICATION   #superfamily hepatitis B virus core antigen
KEYWORDS         core protein
SUMMARY          #length 183 #molecular-weight 21102 #checksum 2199

Query Match      65.6%: Score 1370; DB 2; Length 183;
Best Local Similarity 98.4%: Pred. No. 3,06e-204;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDYKEFGATVELSLSPDSFPFVSVDLDTASALYRALSESPHCSPHNTALROAIL 60
Qy 96 MDIDYKEFGATVELSLSPDSFPFVSVDLDTASALYRALSESPHCSPHNTALROAIL 155
Qy 156 CWGELMTLATWGVNLDDPASRDVLVSVYNTNMGLKFRQLMFIISCLTFGRFVIEYLV 215
Db 61 CWGDLMTLATWGVNLDDPASRDVLVSVYNTNMGLKFRQLMFIISCLTFGRFVIEYLV 120
Db 121 SFGVIRIRPPAYRPPNAIILSLPFTTVYRRGRPRPRRTSPRRRSQSRRRRSSQRE 180
Qy 216 SFGVIRIRPPAYRPPNAIILSLPFTTVYRRGRPRPRRTSPRRRSQSRRRRSSQRE 275
Db 181 SOC 183
Qy 276 SOC 278

RESULT 7
ENTRY          S53216 #type complete
TITLE           e antigen precursor / core antigen - hepatitis B virus
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS        core antigen; e antigen
ORGANISM        #formal_name hepatitis B virus; HBV
VARIETY         #isolate patient Castag/3
DATE            08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
                08-Sep-1997

ACCESSIONS      S53216
REFERENCE        S53112
#authors         Lai, M.E.; Mazzeoli, A.P.; Porru, A.; Balestrieri, A.
#submission      submitted to the EMBL Data Library, March 1995
#accession       S53216
##molecule_type DNA
##residues       1-212 ##label LAI
##cross-references EMBL:X85293; NID:g736124; PID:g736126
##experimental_source isolate patient Castag/3

GENETICS
#gene            C
CLASSIFICATION   #superfamily hepatitis B virus core antigen
KEYWORDS         alternative initiators; core protein
FEATURE
1-29             #domain signal sequence #status predicted #label SIG\
30-178           #product core antigen #status predicted #label CAG\
179-212          #domain carboxyl-terminal propeptide #link ENG #status
                  predicted #label ECP
SUMMARY          #length 212 #molecular-weight 24363 #checksum 752

Query Match      65.6%: Score 1370; DB 2; Length 212;
Best Local Similarity 98.9%: Pred. No. 3,06e-204;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Db	29	GMDIDPYKEFGATVLLSFLSDSPFSRRDLDNLSALYRALSEPEHCSPHHTLRAOI	88
Oy	95	SMDIDPYKEFGATVLLSFLSDSPFSRRDLDNLSALYRALSEPEHCSPHHTLRAOI	154
Db	89	LCMGELMLTLATWGVNLDDPASRDLYSVYNTNMGLKFRQLMFIISCLTIGRETIVIEL	148
Oy	155	LCMGELMLTLATWGVNLDDPASRDLYSVYNTNMGLKFRQLMFIISCLTIGRETIVIEL	214
Db	149	VSGFWIIRTPPAYRPNPAPILISTLETTVVRRRGSSPRRRTPPSPRRRSQSPRRRSOSR	208
Oy	215	VSGFWIIRTPPAYRPNPAPILISTLETTVVRRRGSSPRRRTPPSPRRRSQSPRRRSOSR	274
Db	209	ESQC 212	
Oy	275	ESQC 278	
<hr/>			
RESULT	8		
ENTRY		S53272	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus (isolate patient Licheri-3'90)	
ALTERNATE_NAMES		HBe antigen precursor / Hbc antigen: pre-C/C antigen	
CONTAINS		core antigen; e antigen	
ORGANISM		#formal_name hepatitis B virus, HBV	
DATE		isolate patient Licheri-3'90 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		S53272	
REFERENCE		S53112	
#authors		Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.	
#submitter		submitted to the EMBL Data Library, March 1995	
#accession		S53272	
##molecule_type	DNA		
##residues	1-212	#label LAI	
##cross_references	EMBL:X85315; NID:g736205; PID:g736207		
##experimental_source	isolate patient Licheri-3'90		
GENETICS			
CLASSIFICATION	C	#superfamily hepatitis B virus core antigen	
KEYWORDS		alternative initiators; core protein	
FEATURE			
1-29		#domain signal sequence #status predicted #label SIG\	
30-212		#product core antigen #status predicted #label CAG\	
30-178		#product e antigen #status predicted #label EAG\	
179-212		#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY		#length 212 #molecular_weight 24366 #checksum 446	
Query Match	65.6%; Score 1370; DB 2; Length 212;		
Best Local Similarity	98.9%; Pred. No. 3,06e-204;		
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0			
Db	29	GMDIDPYKEFGATVLLSFLSDSPFSRRDLDNLSALYRALSEPEHCSPHHTLRAOI	88
Oy	95	SMDIDPYKEFGATVLLSFLSDSPFSRRDLDNLSALYRALSEPEHCSPHHTLRAOI	154
Db	89	LCMGELMLTLATWGVNLDDPASRDLYSVYNTNMGLKFRQLMFIISCLTIGRETIVIEL	148
Oy	155	LCMGELMLTLATWGVNLDDPASRDLYSVYNTNMGLKFRQLMFIISCLTIGRETIVIEL	214
Db	149	VSGFWIIRTPPAYRPNPAPILISTLETTVVRRRGSSPRRRTPPSPRRRSQSPRRRSOSR	208
Oy	215	VSGFWIIRTPPAYRPNPAPILISTLETTVVRRRGSSPRRRTPPSPRRRSQSPRRRSOSR	274
Db	209	ESQC 212	
Oy	275	ESQC 278	
<hr/>			
ENTRY	9	S53163	#type complete
TITLE		e antigen precursor / core antigen - hepatitis B virus	

ALTERNATE_NAMES (isolate patient Vitorina'92)
CONTAINS HBe antigen precursor / HBeC antigen: pre-C/C antigen
ORGANISM core antigen: e antigen
#formal_name hepatitis B virus, HBV
#variety isolate patient Vitorina'92
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53163
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Portu, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53163
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85256; NID:q736050; PID:q736052
#experimental_source isolate patient Vitorina'92

GENETICS C
#gene
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label EAG\
#domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24322 #checksum 842

Query Match 65.6%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 4.53e-204;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 88
:|||||
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 154
|||||

Db 89 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 148
|||||

QY 155 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 214
|||||

Db 149 VSFGWIRTPPAYPPNAPILSTPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
|||||

QY 215 VSFGWIRTPPAYPPNAPILSTPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 274
|||||

Db 209 ESOC 212
|||||

QY 275 ESOC 278

RESULT 10
ENTRY S20746 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen: pre-C/C antigen
CONTAINS core antigen: e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient C
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997

ACCESSIONS S20746
REFERENCE S20745
#authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.;
Portu, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with
HBsAg negative chronic liver disease.
#accession S20746
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X65257; NID:q59429; PID:q59431
#experimental_source subtype ayw, patient C

GENETICS C
#gene

CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label EAG\
#domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 506

Query Match 65.6%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4.53e-204;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 88
:|||||
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 154
|||||

Db 89 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 148
|||||

QY 155 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 214
|||||

Db 149 VSFGWIRTPPAYPPNAPILSTPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
|||||

QY 215 VSFGWIRTPPAYPPNAPILSTPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 274
|||||

Db 209 ESOC 212
|||||

QY 275 ESOC 278

RESULT 11
ENTRY S53225 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Chigline-2'86)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen: pre-C/C antigen
CONTAINS core antigen: e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Chigline-2'86
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53225
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Portu, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53225
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85296; NID:q736137; PID:q736139
#experimental_source isolate patient Chigline-2'86

GENETICS C
#gene
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators: core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product core antigen #status predicted #label CAG\
179-212 #product e antigen #status predicted #label EAG\
#domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24308 #checksum 655

Query Match 65.6%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 4.53e-204;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 88
:|||||
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREDALESPPHCSPHNTALRQAI 154
|||||

Db 89 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 148
|||||

QY 155 LCGELMTLATWGVNEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRVTEYL 214
|||||

Db	149	VSFGWITTPPAYRPPNAPILSTLPETTVRRGRGSPRRRPSPPRRRSQSR	208
Oy	215	VSFGWITTPPAYRPPNAPILSTLPETTVRRGRGSPRRRPSPPRRRSQSR	274
Db	209	ESOC 212	
Oy	275	ESOC 278	
RESULT ENTRY	12		
TITLE		553198 #type complete	
ALTERNATE_NAMES		e antigen precursor / core antigen - hepatitis B virus (isolate patient Ferracuti-2'90)	
CONTAINS		HBE antigen precursor / HBC antigen; pre-C/C antigen	
ORGANISM		core antigen; e antigen	
DATE		#formal_name hepatitis B virus, HBV isolate patient Ferracuti-2'90 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		553198	
REFERENCE		553112	
#authors		lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.	
#submission		submitted to the EMBL Data Library, March 1995	
#accession		553198	
##molecule_type		DNA	
##residues		1-212 #label LAI	
##cross-references		EMBL:X85285; NID:g736099; PID:g736100	
##experimental_source		isolate patient Ferracuti-2'90	
GENETICS			
#gene		C	
CLASSIFICATION		#superfamily hepatitis B virus core antigen	
KEYWORDS		alternative initiators; core protein	
FEATURE			
1-29		#domain signal sequence #status predicted #label SIG	
30-212		#product core antigen #status predicted #label CAG\	
30-178		#product e antigen #status predicted #label EAG\	
179-212		#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY		#length 212 #molecular-weight 24292 #checksum 593	
Query Match		65.5%; Score 1368; DB 2; Length 212;	
Best Local Similarity		98.4%; Pred. No. 6, 72e-204;	
Matches 181; Conservative		2; Mismatches 1; Indels 0; Gaps 0	
Db	29	GMDIDPYKEFGATVLLSFLPSAFPSVRDLDTASALYREALSEPHGHALRAOI	88
Oy	95	SMDDIPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSEPHGHALRAOI	154
Db	89	LCWGDMLTLATWGVNLEDPASRDIVSVYNTNMGKFRQLLWFHISCLTFGRETVEYL	148
Oy	155	LCWGLMLTLATWGVNLEDPASRDIVSVYNTNMGKFRQLLWFHISCLTFGRETVEYL	214
Db	149	VSFGWITTPPAYRPPNAPILSTLPETTVRRGRGSPRRRPSPPRRRSQSR	208
Oy	215	VSFGWITTPPAYRPPNAPILSTLPETTVRRGRGSPRRRPSPPRRRSQSR	274
Db	209	ESOC 212	
Oy	275	ESOC 278	
RESULT ENTRY	13		
TITLE		553223 #type complete	
ALTERNATE_NAMES		e antigen precursor / core antigen - hepatitis B virus (isolate patient Chigline-1'85)	
CONTAINS		HBE antigen precursor / HBC antigen; pre-C/C antigen	
ORGANISM		core antigen; e antigen	
DATE		#formal_name hepatitis B virus, HBV isolate patient Chigline-1'85 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS		553223	

REFERENCE	S53112	
#authors	Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.	
#submission	submitted to the EMBL Data Library, March 1995	
#accession	S53223	
#molecule_type	DNA	
#residues	1-212	#label LAI
#cross-references	EMBL:X85295; NID:g736134; PID:g736136	
#experimental_source	isolate patient Chighine-1'85	
GENETICS		
#gene	C	
CLASSIFICATION	#superfamily hepatitis B virus core antigen	
KEYWORDS	alternative initiators; core protein	
FEATURE		
1-29		
30-212	#domain signal sequence #status predicted #label SIG\	
30-178	#product core antigen #status predicted #label CAG\	
179-212	#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY	#length 212 #molecular-weight 24364 #checksum 1123	
Query Match	65.4%; Score 1366; DB 2; Length 212;	
Best Local Similarity	97.8%; Pred. No. 1,48e-203;	
Matches 180; Conservative	3; Mismatches 1; Indels 0; Gaps 0	
Db		
29	GMDIDPKFEGATVLLFFLPDSFPYSVRLDITFASALYREALSPHCSPHHTLROAI	88
95	SMDDIDPKFEGATVLLFFLPDSFPYSVRLDITFASALYREALSPHCSPHHTLROAI	154
89	LCWGLMSLATWGVNLEDPISRDLYSVYVNTNMGKFRQLMFIISCLTFGRTVIEYL	148
155	LCWGLMSLATWGVNLEDPASRDLYSVYVNTNMGKFRQLMFIISCLTFGRTVIEYL	214
Db		
149	VSGFWITPPAPYRPNPILISTPETVYVRGRGSPRRTPSPRRRSQSRRRSQSR	208
Qy	215 VSGFWITPPAPYRPNPILISTPETVYVRGRGSPRRTPSPRRRSQSRRRSQSR	274
Db		
209	ESQC 212	
275	ESQC 278	
RESULT	14	
ENTRY	S53274	#type complete
TITLE	e antigen precursor / core antigen - hepatitis B virus	
ALTERNATE_NAMES	(isolate patient Giordo'84)	
CONTAINS	HBe antigen precursor / HBe antigen; pre-C/C antigen	
ORGANISM	core antigen; e antigen	
#variety	#formal_name hepatitis B virus, HBV	
DATE	isolate patient Giordo'84	
	08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997	
ACCESSIONS	S53274	
REFERENCE	S53112	
#authors	Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.	
#submission	submitted to the EMBL Data Library, March 1995	
#accession	S53274	
#molecule_type	DNA	
#residues	1-212	#label LAI
#cross-references	EMBL:X85316; NID:g736208; PID:g736210	
#experimental_source	isolate patient patient Giordo'84	
GENETICS		
#gene	C	
CLASSIFICATION	#superfamily hepatitis B virus core antigen	
KEYWORDS	alternative initiators; core protein	
FEATURE		
1-29		
30-212	#domain signal sequence #status predicted #label SIG\	
30-178	#product core antigen #status predicted #label CAG\	
179-212	#domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP	
SUMMARY	#length 212 #molecular-weight 24376 #checksum 390	
Query Match	65.4%; Score 1366; DB 2; Length 212;	

Best Local Similarity 98.98; Pred. No. 1.48e-203;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEGATVEELSTFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
QY 95 SMDIDPYKEGATVEELSTFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 154
Db 89 LCMGELMTLATWGVNLEDPAASRDVSVYNTNMGKFRQLMFHISCLTFGRETVEIYL 148
QY 155 LCMGELMTLATWGVNLEDPAASRDVSVYNTNMGKFRQLMFHISCLTFGRETVEIYL 214
Db 149 VSGVWIRTPPAYRPPNAPILITLPTTVYRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 215 VSGVWIRTPPAYRPPNAPILITLPTTVYRRGRSPRRRTSPRRRSQSPRRRSQSR 274
Db 209 ESOC 212
QY 275 ESOC 278

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu'89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Muresu'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53169
REFERENCE S53112
#authors Lai, M.E.; Mazzeoni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
##molecule_type DNA
##residues 1-183 #label LAI
##cross-references EMBL:X85275; NID:q736057; PID:q736060
##experimental_source isolate patient Muresu'89
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS C
#gene #superfamily hepatitis B virus core antigen
CLASSIFICATION core protein
KEYWORDS #length 183 #molecular_weight 21088 #checksum 1849
SUMMARY

Query Match 65.3%; Score 1364; DB 2; Length 183;
Best Local Similarity 98.48; Pred. No. 3.26e-203;
Matches 180; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPYKEGATVEELSTFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
QY 96 MDIDPYKEGATVEELSTFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 155
Db 61 CWMGELMTLATWGVNLEDPAASRDVSVYNTNMGKFRQLMFHISCLTFGRETVEIYL 120
QY 156 CWMGELMTLATWGVNLEDPAASRDVSVYNTNMGKFRQLMFHISCLTFGRETVEIYL 215
Db 121 SFGVWIRTPPAYRPPNAPILITLPTTVYRRGRSPRRRTSPRRRSQSPRRRSQSR 180
QY 216 SFGVWIRTPPAYRPPNAPILITLPTTVYRRGRSPRRRTSPRRRSQSPRRRSQSR 275
Db 181 SOC 183
QY 276 SOC 278

Search completed: Thu Dec 16 13:24:56 1999
Job time : 57 secs.

WISCONSIN
(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSPRCH-PP protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:25:14 1999; MasPar time 9.63 seconds
Tabular output not generated. 816.384 Million cell updates/sec

Title: >US3869-29-38183
Description: (1-278) from us3869-29-38183.pep
Perfect Score: 2088
Sequence: 1 MDIDPYKFGATVELLSFLP.....RRRSQSPRRRSQSPRESQC 278

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 48.215; Variance 102.691; scale 0.470

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1374	65.8	183	1	CORA_HPBVA CORE ANTIGEN.	9.32e-236
2	1358	65.0	211	1	CORA_HPBVA CORE ANTIGEN.	1.32e-232
3	1349	64.6	183	1	CORA_HPBVA CORE ANTIGEN.	7.85e-231
4	1336	64.0	183	1	CORA_HPBVA CORE ANTIGEN.	2.85e-228
5	1334	63.9	183	1	CORA_HPBVA CORE ANTIGEN.	7.06e-228
6	1332	63.3	183	1	CORA_HPBVA CORE ANTIGEN.	1.62e-225
7	1331	63.3	183	1	CORA_HPBVA CORE ANTIGEN.	2.56e-225
8	1314	62.9	195	1	CORA_HPBVA CORE ANTIGEN.	6.10e-224
9	1313	62.9	212	1	CORA_HPBVA CORE ANTIGEN.	9.59e-224
10	1308	62.6	183	1	CORA_HPBVA CORE ANTIGEN.	9.24e-223
11	1305	62.5	185	1	CORA_HPBVA CORE ANTIGEN.	3.60e-222
12	1304	62.5	214	1	CORA_HPBVA CORE ANTIGEN.	5.66e-222
13	938	44.9	188	1	CORA_HPBVA CORE ANTIGEN.	1.96e-150
14	933	44.7	187	1	CORA_HPBVA CORE ANTIGEN.	1.83e-149
15	934	44.7	217	1	CORA_HPBVA CORE ANTIGEN.	1.17e-149
16	208	10.0	493	1	CETP_HUMAN CHOLESTERYL ESTER TRAN	1.19e-15
17	208	10.0	493	1	CETP_HUMAN CHOLESTERYL ESTER TRAN	1.19e-15
18	205	9.8	389	1	CETP_HUMAN CHOLESTERYL ESTER TRAN	3.58e-15
19	200	9.6	305	1	CORA_HPBVA CORE ANTIGEN.	2.21e-14
20	200	9.6	305	1	CORA_HPBVA CORE ANTIGEN.	2.21e-14
21	200	9.6	305	1	CORA_HPBVA CORE ANTIGEN.	2.21e-14
22	200	9.6	305	1	CORA_HPBVA CORE ANTIGEN.	2.21e-14
23	200	9.6	305	1	CORA_HPBVA CORE ANTIGEN.	2.21e-14

RESULT	ID	SCORE	STANDARD	PROT	ALIGNMENTS
24	171	8.2	497	1	CETP_RABIT CHOLESTERYL ESTER TRAN
25	158	7.6	196	1	SPR2_CAEEL PUTATIVE SPLICING FACT
26	148	7.1	208	1	YSX2_CAEEL HYPOHETICAL 24.0 KD P
27	136	6.5	78	1	PR11_SEPOF SPERMATID-SPECIFIC PRO
28	131	6.3	61	1	SPR1_MACCU SPERM PROTAINE P1
29	130	6.2	77	1	PR12_SEPOF SPERMATID-SPECIFIC PRO
30	125	6.0	60	1	SPR1_MACGI SPERM PROTAINE P1
31	125	6.0	498	1	SPR1_MACRO SPERM PROTAINE P1
32	126	6.0	498	1	VE2_HPV08 REGULATORY PROTEIN E2
33	126	6.0	1523	1	SON_HUMAN SON PROTEIN (SON3)
34	124	5.9	57	1	HSP1_DIDNA SPERM PROTAINE P1
35	121	5.8	61	1	HSP1_CHICK SPERM HISTONE (PROTAMI
36	122	5.8	91	1	PH11_MYTED SPERM-SPECIFIC PROTEIN
37	122	5.8	132	1	PR1_AMTGR PUTATIVE PRE-MRNA SPLI
38	121	5.8	739	1	DD13_CAEEL PUTATIVE PRE-MRNA SPLI
39	121	5.8	843	1	CYP1_BRUMA PEPTIDYLPROLYL ISOMERA
40	119	5.7	50	1	HSP1_MAT SPERM PROTAINE P1 (CY
41	119	5.7	50	1	HSP1_MOUSE SPERM PROTAINE P1 (CY
42	118	5.7	238	1	SPR7_HUMAN SPERMATID-SPECIFIC PRO
43	118	5.7	576	1	CAT1_RHOCA PEROXIDASE / CATALASE
44	115	5.5	102	1	HSP2_MACMG SPERM HISTONE P2 PRECU
45	115	5.5	104	1	HSP2_CALUA SPERM HISTONE P2 PRECU

US3869-29-38183.rsp

Query Match: 65.8%; Score 1374; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 9.32e-236;

Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 MDIPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALROAIL 60
  |||
Qy 96 MDIPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALROAIL 155
  |||
Db 61 CWGELMTLATWGVNLEDPASRDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 120
  |||
Qy 156 CWGELMTLATWGVNLEDPASRDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 215
  |||
Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQRE 180
  |||
Qy 216 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQRE 275
  |||
Db 181 SOC 183
  |||
Qy 276 SOC 278
```

RESULT 2
ID CORA_HPBYA STANDARD; PRT; 211 AA.

AC P24023;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.

OS HEPATITIS B VIRUS (STRAIN ALPHA1).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
RP SEQUENCE FROM N.A.

RA TONG S., LI J., VIRVITSKI L., TREPO C.;
RT "Active hepatitis B virus replication in the presence of anti-HBe is
associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M32138; G495034; ALT_SEQ.

DR PIR; A34773; NKVLAL.

DR PFAM; PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT.

FT REPEAT 183 190

FT REPEAT 198 206

SEQUENCE 211 AA; 24208 MW; 0924CE5E CRC32;

Query Match 65.0%; Score 1358; DB 1; Length 211;

Best Local Similarity 95.2%; Pred. No. 1,32e-232;

Matches 179; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Db 24 LGWIDMDIDPKFEGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTAL 83
  |||
Qy 91 LGWIDMDIDPKFEGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTAL 150
  |||
Db 84 ROAILCWDGLTLTSTWGVNLEDPDPTSDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 143
  |||
Qy 151 ROAILCWDGLTLTSTWGVNLEDPDPTSDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 210
  |||
Db 144 IEYLVSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRR 203
  |||
Qy 211 IEYLVSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRR 270
  |||
Db 204 TQSRSSOC 211
  |||
Qy 271 TQSRSSOC 278
```

RESULT 3
ID CORA_HPBYZ STANDARD; PRT; 183 AA.

AC P0147;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.

OS HEPATITIS B VIRUS (SUBTYPE ADYN).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
RP SEQUENCE FROM N.A.

RA PASER M., GOTO T., GILBERT W., ZINK B., SCHALLER H., MCKAY P.,

RA LEADBETTER G., MURRAY K.;

RT "Hepatitis B virus genes and their expression in E. coli.";

RL NATURE 282:575-579(1979).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; J02202; G329638; -

DR EMBL; A08967; G411874; -

DR PIR; B93217; NKVLAL.

DR PFAM; PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT.

FT REPEAT 162 169

FT REPEAT 170 177

SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;

Query Match 64.6%; Score 1349; DB 1; Length 183;

Best Local Similarity 96.7%; Pred. No. 7,85e-231;

Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```
Db 1 MDIPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALROAIL 60
  |||
Qy 96 MDIPYKEFGATVLLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALROAIL 155
  |||
Db 61 CWGELMTLATWGVNLEDPASRDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 120
  |||
Qy 156 CWGELMTLATWGVNLEDPASRDVLVS YVNTNMGKFRQLLMFHISCLTFGRETVEYLV 215
  |||
Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQRE 180
  |||
Qy 216 SFGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQRE 275
  |||
Db 181 SOC 183
  |||
Qy 276 SOC 278
```

RESULT 4
ID CORA_HPBYO STANDARD; PRT; 183 AA.

AC P17392;
DT 01-AUG-1990 (REL. 15, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN OKINAWA/POOW282).

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
RP SEQUENCE FROM N.A.

RA MEDLINE; 89010694.

RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASUROSEWIGNO R.I., IMAI M.,

RA MIYAKAWA Y., MATSUMI K.;

RT "Typing hepatitis B virus by homology in nucleotide sequence:
comparison of surface antigen subtypes.";

[illegible]

```

RESULT      7
ID          ID CORA_HPBW STANDARD; PRT: 185 AA.
AC          P03149;
DT          21-JUL-1986 (REL. 01, CREATED)
DT          21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT          01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE          CORE ANTIGEN.
CC          C.
OS          HEPATITIS B VIRUS (SUBTYPE ADW).
OC          VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN          [1]
RP          SEQUENCE FROM N.A.
RX          MEDLINE: 83168919.
RA          ONO Y., ONDA H., ICARASHI K., SUGINO Y., NISHIOKA K.:
RT          "The complete nucleotide sequences of the cloned hepatitis B virus
RL          DNA: subtype adr and adw."
RL          NUCLEIC ACIDS RES. 11:1747-1757(1983).
CC          -----
CC          CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          CC the European Bioinformatics Institute. There are no restrictions on
CC          CC use by non-profit institutions as long as its content is in no way
CC          CC modified and this statement is not removed. Usage by and for commercial
CC          CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC          CC or send an email to license@isb-sib.ch).

```

[illegible]

Db	121	SEGVWITTPPAYRRPNNAPIILSTLPEETVVRARDGRSPRRRTSPRRRSQSPRRRSQS	180
QY	216	SEGVWITTPPAYRRPNNAPIILSTLPEETVVRARR--GRSPRRRTSPRRRSQSPRRRSQS	273
Db	181	RESEC	185
QY	274	RESEC	278
RESULT	8	STANDARD;	PRT; 195 AA.
ID	CORA_HPBYF		
AC	P29178;		
DT	01-DEC-1992 (REL. 24, CREATED)		
DT	01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)		
DT	01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)		
DE	CORE ANTIGEN.		
GN	C.		
OS	HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF).		
OC	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE; 90169850.		
RT	BHART R.A., ULRICH P.P., VYAS G.N.;		
RT	"Molecular characterization of a new variant of hepatitis B virus in		
RL	a persistently infected homosexual man.";		
RL	HEPATOLOGY 11:271-276(1990).		
DR	P1R; A37182; NIKVH3.		
DR	PFAM: PF00906; Hepatitis_core; 1.		
KW	CORE PROTEIN; REPEAT		
FT	REPEAT	174	181
FT	REPEAT	182	189
QO	SEQUENCE	195 AA; 22461 MW; AF3DB5F3	CRC32;

Query Match	62.9%	Score 1314	DB 1	Length 195
Best Local Similarity	93.5%	Pred. No. 6.10e-224		
Matches 173	Conservative	9	Mismatches 3	Indels 0
			Gaps	0

Db	11	FGLDIDPYKEGATVELLSFLPSDFPSPVDLLDTASALYRESLESDDHCSPHHTALRQA	70
		:::	
Qy	94	LSMDIDPYKEGATVELLSFLPSDFPSPVDLLDTASALYRESLESDDHCSPHHTALRQA	153
Db	71	ILCVELMLTATWGNLLEDPASRDVYVNTNMGKIKIROLAMFHNSCTFGRETVLX	130
Qy	154	ILCVELMLTATWGNLLEDPASRDVYVNTNMGKIKIROLAMFHNSCTFGRETVLX	213
Db	131	LVSGVWIKTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRRSSQS	190
Qy	214	LVSGVWIKTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTSPRRRRSSQS	273
Db	191	RESOC 195	
Qy	274	RESOC 278	

RESULT	9	STANDARD;	PRT;	212 AA.
ID	CORA_HPBVT			
AC	005495;			
DT	01-NOV-1995 (REL. 32, CREATED)			
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			
OS	HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE M4B).			
OC	VIRUSSES; RETROID VIRUSES; HERPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEULANE; 93346970.			
RA	NUMMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,			
RA	GERLICH W.H.;			
RT	"identification of a new hepatitis B virus (HBV) genotype from Brazil			
RT	that expresses HBV surface antigen subtype adw4.";			
RL	J. GEN. VIROL. 74:1627-1632(1993).			
CC	-----			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X69798: G59423: -
DR PFAM: PF00906: Hepatitis_core: 1.
KM CORE PROTEIN: REPEAT.
FT DOMAIN 178 204 ARG-RICH.
FT DOMAIN 184 203 3 X 5 AA REPEATS OF S-P-R-R-R.
FT REPEAT 184 188 1.
FT REPEAT 191 195 2.
FT REPEAT 199 203 3.
SQ SEQUENCE 212 AA: 24234 MW: BCCDF263 CRC32:

Query Match 62.9%; Score 1313; DB 1; Length 212;
Best Local Similarity 92.9%; Pred. No. 9.59e-224;
Matches 171; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

Db 29 GMDIPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
OY 95 SMDIPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 154
Db 89 LCMGELMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 148
OY 155 LCMGELMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 214
Db 149 VSEGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSP 208
OY 215 VSEGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSP 274
Db 209 ASQC 212
OY 275 ESQC 278

RESULT 10
ID CORA_HPBVL STANDARD; PRT; 183 AA.
AC P12901;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN LSH / CHIMPANZEE ISOLATE).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88258473.
RA VAUDIN M., WOLSTENHOLME A.J., TSIQUAYE K.N., ZUCKERMAN A.J.,
RA HARRISON T.J.;
RT "The complete nucleotide sequence of the genome of a hepatitis B
RT virus isolated from a naturally infected chimpanzee.";
RL J. GEN. VIROL. 69:1383-1389(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00220: D100603: -
DR PIR: A28885: NKVLCP.
DR PFAM: PF00906: Hepatitis_core: 1.
KM CORE PROTEIN: REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 20999 MW: 1AF57C9 CRC32:

Query Match 62.6%; Score 1308; DB 1; Length 183;
Best Local Similarity 94.5%; Pred. No. 9.24e-223;
Matches 173; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 60
OY 96 MDIDPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 155
Db 61 CWSGLMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 120
OY 156 CWSGLMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 215
Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSPA 180
OY 216 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSPRE 275
Db 181 SOC 183
OY 276 SOC 278

RESULT 11
ID CORA_HPBV2 STANDARD; PRT; 185 AA.
AC P03148;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW2).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA VALENZUELA P., QUIROGA M., ZALDIVAR J., GRAY P., RUTTER W.J.;
RL (IN) FIELD B.N., JAKENISCH R., FOX C.F. (EDS.);
RL ANIMAL VIRUS GENETICS, PP.57-70, ACADEMIC PRESS, NEW YORK (1980).
DR PIR: A94409; NKVLA3.
DR PFAM: PF00906: Hepatitis_core: 1.
KM CORE PROTEIN: REPEAT.
FT REPEAT 164 171
FT REPEAT 172 179
SQ SEQUENCE 185 AA: E2EA3360 CRC32:

Query Match 62.5%; Score 1305; DB 1; Length 185;
Best Local Similarity 95.7%; Pred. No. 3.60e-222;
Matches 177; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Db 1 MDIDPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 60
OY 96 MDIDPYKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 155
Db 61 CWSGLMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 120
OY 156 CWSGLMTLAWGNLDDPSRDLYVNTNMGKIRQLMFIISCLTGRETVLEYL 215
Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSP 180
OY 216 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSP 273
Db 181 RESQC 185
OY 274 RESQC 278

RESULT 12
ID CORA_HPBV9 STANDARD; PRT; 214 AA.
AC P17099;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.

```

OS HEPATITIS B VIRUS (SUBTYPE ADM / STRAIN 991).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA KOEHEL H.G., SCHUELER A., LOTTWANN S., THOMSEN R.;
RC SUBMITTED (FEB-1990) TO EMBL/GENBANK/DBJ DATA BANS.
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: X51970; G60433; -.
DR PIR: S10381; NKVLKS.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 193 200
FT REPEAT 201 208
SQ SEQUENCE 214 AA; 24722 MW; 786CD048 CRC32;

Query Match 62.5%; Score 1304; DB 1; Length 214;
Best Local Similarity 95.2%; Pred. No. 5.66e-222;
Matches 177; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

Db 29 GMDIDPKKEGATVELLSFSPDFSVRLDLPASALYREALSPHCSPHHTALROAI 88
QY 95 SMIDPKKEGATVELLSFSPDFSVRLDLPASALYREALSPHCSPHHTALROAI 154
Db 89 LCMGELMTLATWGVNLEDPASRDLYVYNTNGKIRQLMFRISYLTGEGTVEYL 148
QY 155 LCMGELMTLATWGVNLEDPASRDLYVYNTNGKIRQLMFRISYLTGEGTVEYL 214
Db 149 VSGVWIRTPAPYRPPNAPILSTLPETTVRRDRGSPRRRPSRRRSQSPRRRSQ 208
QY 215 VSGVWIRTPAPYRPPNAPILSTLPETTVRRR--GRSPRRTPSPRRRSQSPRRRSQ 272
Db 209 SRESOC 214
QY 273 SRESOC 278

RESULT 13
ID CORA_MHV1 STANDARD; PRT; 188 AA.
AC P03152;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 1 (MHV 1),
OS WOODCHUCK HEPATITIS VIRUS 7 (MHV 7),
OS WOODCHUCK HEPATITIS VIRUS 59 (MHV 59), AND
OS WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE) (MHV 8).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA GALBERT F., CHEN T.N., MANDART E.;
RX MEDLINE: 82216969.
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT comparison with the hepatitis B virus sequence."
RL J. VIROL. 41:51-65(1982).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-WOODCHUCK HEPATITIS VIRUS 1;
RX MEDLINE: 82216969.
RA COHEN J.I., MILLER R.H., ROSENBLUM B., DENNISTON K., GERIN J.L.,
RA PURCELL R.H.;
RT "Sequence comparison of woodchuck hepatitis virus replicative forms
RT shows conservation of the genome."

```

```

RL VIROLOGY 162:12-20(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE);
RX MEDLINE: 89184524.
RA GIRONES R., COTE P.J., HORNBUCKLE W.E., TENNANT B.C., GERIN J.L.,
RA PURCELL R.H., MILLER R.H.;
RT "Complete nucleotide sequence of a molecular clone of woodchuck
RT hepatitis virus that is infectious in the natural host."
RL PROC. NATL. ACAD. SCI. U.S.A. 86:1846-1849(1989).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: J02442; G336129; -.
DR EMBL: M18752; G336140; -.
DR EMBL: M19183; G336145; -.
DR EMBL: J04514; G336149; -.
DR PIR: A03713; NKVLC.
DR PIR: C32397; NKVLC.
DR PIR: C32397; NKVLC4.
DR PIR: C32397; NKVLC3.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 167 174
FT REPEAT 175 182
SQ SEQUENCE 188 AA; 21693 MW; A667DB27 CRC32;

Query Match 44.9%; Score 938; DB 1; Length 188;
Best Local Similarity 67.0%; Pred. No. 1.96e-150;
Matches 126; Conservative 25; Mismatches 32; Indels 5; Gaps 1;

Db 1 MDIDPKKEGSSYQLNPLDFFPDNLAVDTATALYEEELTGREHCSPHHTALROAI 60
QY 96 MDIDPKKEGATVELLSFSPDFSVRLDLPASALYREALSPHCSPHHTALROAI 155
Db 61 CWDELTKLIAMSSNITSQVRIIYNHNDTWGLKVRQSLMFHLSCLTFGCHTQOEFLV 120
QY 156 CWGELMTLATWGVNLEDPASRDLYVYNTNGKIRQLMFRISYLTGEGTVEYL 215
Db 121 SFGVWIRTPAPYRPPNAPILSTLPETTVRRRGARASPPRRRTPSPRRRSQSPRRR 180
QY 216 SFGVWIRTPAPYRPPNAPILSTLPETTVRRG-----RSPARRTPSPRRRSQSPRRR 270
Db 181 SOSPSANC 188
QY 271 SOSRESOC 278

RESULT 14
ID CORA_MHV8 STANDARD; PRT; 187 AA.
AC P06433;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 8 (MHV 8).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA KODAMA K., OGASAWARA N., YOSHIKAWA H., MORAKAMI S.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT evolutionary relationship between hepadnaviruses."
RL J. VIROL. 56:978-986(1985).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

 W P E R E H
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

msrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:26:11 1999; Maspar time 19.81 Seconds

Tabular output not generated. 765,940 Million cell updates/sec

Title: >US3869-29-38183
 Description: (1-278) from US3869-29-38183.pep
 Perfect score: 2088
 Sequence: 1 MDIDPYKEFGATVELLSFLP.....RRRRSQSPRRRRSQSRRESQC 278

Scoring table:
 PAM 150
 Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

splemb19
 1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 46.475; Variance 103.566; scale 0.449

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	65.9	212 14	089656	PRE-C/CORE.	3.17e-226
2	1373	65.8	183 14	089437	X, PREC AND C GENES (C	7.57e-226
3	1373	65.8	183 14	068008	X, PREC AND C GENES (F	7.57e-226
4	1374	65.8	212 14	068020	PRE-C/CORE.	4.90e-226
5	1374	65.8	212 14	089597	HBCAG.	7.57e-226
6	1373	65.8	212 14	067876	PRE C/C ORF.	4.90e-226
7	1372	65.7	212 14	011884	CORE ANTIGEN PRECURSOR	1.17e-225
8	1370	65.6	183 14	068066	X, PREC AND C GENES (L	2.78e-225
9	1370	65.6	212 14	068066	PRE-C/CORE.	2.78e-225
10	1370	65.6	212 14	068025	PRE-C/CORE.	2.78e-225
11	1369	65.6	212 14	067872	PRE C/C ORF.	4.30e-225
12	1369	65.6	212 14	067984	PRE-C/CORE.	4.30e-225
13	1369	65.6	212 14	068032	PRE-C/CORE.	4.30e-225
14	1368	65.5	212 14	068010	PRE-C/CORE.	6.64e-225
15	1366	65.4	212 14	068070	PRE-C/CORE.	1.58e-224
16	1366	65.4	212 14	068030	PRE-C/CORE.	1.58e-224
17	1364	65.3	183 14	067989	X, PREC AND C GENES (M	3.77e-224
18	1363	65.3	183 14	068048	X, PREC AND C GENES (F	5.82e-224
19	1363	65.3	212 14	068014	PRE-C/CORE.	5.82e-224
20	1363	65.3	212 14	068012	PRE-C/CORE.	5.82e-224

21	1363	65.3	212 14	067980	PRE-C/CORE.	5.82e-224
22	1362	65.2	212 14	068051	PRE-C/CORE.	8.98e-224
23	1360	65.1	212 14	089446	PRE-C/CORE.	2.14e-223
24	1360	65.1	212 14	068016	PRE-C/CORE.	2.14e-223
25	1360	65.1	212 14	096845	COMPLETE GENOME.	2.14e-223
26	1359	65.1	212 14	068053	PRE-C/CORE.	3.31e-223
27	1357	65.0	212 14	068075	PRE-C/CORE.	7.88e-223
28	1357	65.0	212 14	068077	PRE-C/CORE.	7.88e-223
29	1355	64.9	212 14	068045	PRE-C/CORE.	1.88e-222
30	1352	64.8	183 14	089531	CORE PROTEIN.	6.90e-222
31	1351	64.7	183 14	067946	C ANTIGEN.	1.07e-221
32	1351	64.7	212 14	089719	PRE-C/CORE.	1.07e-221
33	1349	64.6	183 14	067997	X, PREC AND C GENES (B	2.54e-221
34	1348	64.6	183 14	068037	X, PREC AND C GENES (D	3.92e-221
35	1348	64.6	183 14	068060	X, PREC AND C GENES (M	3.92e-221
36	1347	64.5	212 14	067912	PRE-CORE/CORE PROTEIN.	6.05e-221
37	1345	64.4	183 14	067973	X, PREC AND C GENES (T	1.44e-220
38	1344	64.4	212 14	092918	PRE-CORE PROTEIN PRECUR	2.22e-220
39	1344	64.4	212 14	068035	PRE-C/CORE.	2.22e-220
40	1343	64.3	183 14	068023	X, PREC AND C GENES (C	3.43e-220
41	1342	64.3	183 14	067964	X, PREC AND C GENES (F	5.30e-220
42	1341	64.2	183 14	068064	X, PREC AND C GENES (L	8.17e-220
43	1341	64.2	183 14	068003	X, PREC AND C GENES (S	8.17e-220
44	1341	64.2	212 14	068042	PRE-C/CORE.	8.17e-220
45	1341	64.2	212 14	081115	HBCAG.	8.17e-220

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	212 AA.
AC	089656;			
DT	01-NOV-1996 (TREMBAEEL. 01, CREATED)			
DT	01-NOV-1996 (TREMBAEEL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1996 (TREMBAEEL. 08, LAST ANNOTATION UPDATE)			
DE	PRE-C/CORE.			
GN	PRE-C/CORE.			
OS	HEPATITIS B VIRUS.			
OC	HEPATITIS B VIRUS; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RN	SEQUENCE FROM N.A.			
RC	STRAIN-PATIENT LICHENI-1/85;			
RA	LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;			
RL	SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	SEQUENCE FROM N.A.			
RP	STRAIN-ATW4;			
RA	PLUCIENNICZAK A.;			
RL	SUBMITTED (AUG-1994) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE: 81012091.			
RA	GALBERT F., MANDART E., FITOUSSI F., TIOUAI P., CHARNAY P.;			
RT	"Nucleotide sequence of the hepatitis B virus genome (subtype ayw)			
RL	cloned in E. coli.;"			
RL	NATURE 281:646-650(1979).			
RN	SEQUENCE FROM N.A.			
RP	SEQUENCE FROM N.A.			
RA	BORISOVA G.P., POMPEY P.P., BYCHKO V.V., PUSHKO P.M., KALIS Y.V.;			
RL	DISILER A.V., GREN E.Y., TSIBINOGIN V.V., KUKAIN R.A.;			
RL	DOKL. BIOCHEM. 279:386-390(1985).			
DR	EMBL: X85290; G736116; -			
DR	EMBL: X85300; G736152; -			
DR	EMBL: X85313; G736196; -			
DR	EMBL: X85283; G736090; -			
DR	EMBL: X85306; G736174; -			
DR	EMBL: X85306; G527437; -			
DR	EMBL: J02203; G329642; -			
DR	EMBL: X85312; G736193; -			
DR	PFAM: PF00906; Hepatitis-core; 1.			
KM	SIGNAL.			
SO	SEQUENCE 212 AA; 24350 MM; 71EA2C82 CRC32;			

Query Match 65.9%; Score 1375; DB 14; Length 212;
 Best Local Similarity 99.5%; Pred. No. 3.17e-226;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 88
 :
 QY 95 SMDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 154
 :
 Db 89 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 148
 :
 QY 155 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 214
 :
 Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 208
 :
 QY 215 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 274
 :
 Db 209 ESQC 212
 :
 QY 275 ESQC 278

RESULT 2 PRELIMINARY; PRT; 183 AA.

ID 089437
 AC 089437;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE X. PREC AND C GENES (CASTAA 2).
 GN CORE.
 OS HEPATITIS B VIRUS.
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT CASTAA-2'87;
 RA LAI M.E., MAZOLENT A.P., PORRU A., BALESTRIERI A.;
 RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AYW;
 RA KARAYIANNIS P.;
 RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AYW;
 RA KARAYIANNIS P.;
 RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: X85289; G736113; -;
 DR EMBL: X80925; E198085; -;
 DR PFM: PF00906; Hepatitis_core: 1;
 SO SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 65.8%; Score 1373; DB 14; Length 183;
 Best Local Similarity 99.5%; Pred. No. 7.57e-226;
 Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 60
 :
 QY 96 MDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 155
 :
 Db 61 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 120
 :
 QY 156 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 215
 :
 Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 180
 :
 QY 216 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 275
 :
 Db 181 SOC 183
 :
 QY 276 SOC 278

RESULT 3

ID 068008 PRELIMINARY; PRT; 183 AA.

AC 068008;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE X. PREC AND C GENES (FERRACUTTI 1).
 GN CORE.
 OS HEPATITIS B VIRUS.
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT FERRACUTTI-1'89;
 RA LAI M.E., MAZOLENT A.P., PORRU A., BALESTRIERI A.;
 RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: X85284; G736098; -;
 DR PFM: PF00906; Hepatitis_core: 1;
 SO SEQUENCE 183 AA; 21102 MW; BB9A7FB3 CRC32;

Query Match 65.8%; Score 1373; DB 14; Length 183;
 Best Local Similarity 99.5%; Pred. No. 7.57e-226;
 Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 60
 :
 QY 96 MDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 155
 :
 Db 61 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 120
 :
 QY 156 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 215
 :
 Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 180
 :
 QY 216 SFGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 275
 :
 Db 181 SOC 183
 :
 QY 276 SOC 278

RESULT 4 PRELIMINARY; PRT; 212 AA.

ID 068020;
 AC 068020;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE PRE-C/CORE.
 GN PRE-C/CORE.
 OS HEPATITIS B VIRUS.
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT CASTAG-1'85;
 RA LAI M.E., MAZOLENT A.P., PORRU A., BALESTRIERI A.;
 RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: X85291; G736119; -;
 DR PFM: PF00906; Hepatitis_core: 1;
 SO SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 65.8%; Score 1374; DB 14; Length 212;
 Best Local Similarity 98.9%; Pred. No. 4.90e-226;
 Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 88
 :
 QY 95 SMDIDPKKEGATVELLSFLPSDFPSPVRDLDTASALYREALSPHCSPHHTALROAI 154
 :
 Db 89 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 148
 :
 QY 155 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETVIEYL 214
 :
 Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 208
 :
 QY 215 VSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 274
 :

Db 209 ESOC 212
OY 275 ESOC 278

RESULT 5
ID 089597 PRELIMINARY; PRT; 212 AA.
AC 089597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HECAG;
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SUB-TYPE AYW;
RX MEDLINE: 94079539.
RA PREISLER-ADAMS S., SCHLAYER M.J., PETERS T., HETTLER F., GEROK W.,
RA RASENACK J.;
RT "Sequence analysis of hepatitis B virus DNA in immunologically
RT negative infection."
RL ARCH. VIROL. 133:385-396(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RN SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X72702; G288930; -;
DR EMBL: X80925; E198084; -;
DR PFAM: PF00906; Hepatitis core: 1
SQ SEQUENCE 212 AA; 24336 MW; 575D20BF CRC32;

Query Match 65.8%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4.90e-226;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
OY 95 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 148
OY 155 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 214

Db 149 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
OY 215 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 274

Db 209 ESOC 212
OY 275 ESOC 278

RESULT 6
ID 067876 PRELIMINARY; PRT; 212 AA.
AC 067876;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CI, HBV SUBTYPE AYW;
RC STRAIN-PATIENT CI, BALESTRIERI A., MELIS A., PORRU A.;
RC LAI M.E., MAZOLENTI A.P., BALESTRIERI A., MELIS A., PORRU A.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65258; G59436; -;
DR PFAM: PF00906; Hepatitis core: 1.
SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;

Query Match 65.8%; Score 1373; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 7.57e-226;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
OY 95 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 148
OY 155 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 214

Db 149 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
OY 215 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 274

Db 209 ESOC 212
OY 275 ESOC 278

RESULT 7
ID 011884 PRELIMINARY; PRT; 212 AA.
AC 011884;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN PRECURSOR.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA RAO B.S., CASEY J.L., RINAUDO J.S., KORBA B.E.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U95551; G2182120; -;
DR PFAM: PF00906; Hepatitis core: 1.
SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;

Query Match 65.7%; Score 1372; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.17e-225;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
OY 95 SMDIDPKKEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 148
OY 155 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMWHISCLTFGRTVIEYL 214

Db 149 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
OY 215 VSEFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 274

Db 209 ESOC 212
OY 275 ESOC 278

RESULT 8
ID 068066 PRELIMINARY; PRT; 183 AA.
AC 068066;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X. PREC AND C GENES (LICHERI 2).
OS CORE.
OS HEPATITIS B VIRUS.

```
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2'87;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 6F38A3B CRC32;

Query Match
Best Local Similarity 98.4%; Score 1370; DB 14; Length 183;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 60
   |||
QY 96 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 155
   |||
Db 61 CWDMLTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 120
   |||
QY 156 CWDMLTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 215
   |||
Db 121 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
   |||
QY 216 SFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 275
   |||
Db 181 SQC 183
   |||
QY 276 SQC 278

RESULT 9
ID 068068; PRELIMINARY; PRT; 212 AA.
AC 068068;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3'90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
   |||
QY 95 SMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 154
   |||
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHICCLFGRETVEYLV 148
   |||
QY 155 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 214
   |||
Db 149 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
   |||
QY 215 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
   |||
Db 209 ESQC 212
   |||
QY 275 ESQC 278

RESULT 10
ID 068025; PRELIMINARY; PRT; 212 AA.
```

```
AC 068025;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG'3;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;

Query Match
Best Local Similarity 98.9%; Score 1370; DB 14; Length 212;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
   |||
QY 95 SMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 154
   |||
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 148
   |||
QY 155 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 214
   |||
Db 149 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
   |||
QY 215 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
   |||
Db 209 ESQC 212
   |||
QY 275 ESQC 278

RESULT 11
ID 067872; PRELIMINARY; PRT; 212 AA.
AC 067872;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C, HBSAG POSITIV, HBV SUBTYPE ATW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -.
PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24348 MW; A6BD9DA1 CRC32;

Query Match
Best Local Similarity 98.4%; Score 1369; DB 14; Length 212;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
   |||
QY 95 SMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 154
   |||
Db 89 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 148
   |||
QY 155 LCGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGFRETVEYLV 214
   |||
Db 149 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
   |||
QY 215 VSRGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 274
   |||
Db 209 ESQC 212
```

1111
QY 275 ESOC 278

RESULT 12
ID 067984 PRELIMINARY; PRT: 212 AA.
AC 067984;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT VITTORINA'92;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85256; G736052; -
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24322 MW; 83D9780B CRC32.

Query Match 65.6%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4,30e-225;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 88
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 148
QY 155 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 214

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
QY 215 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 274

Db 209 ESOC 212
QY 275 ESOC 278

RESULT 13
ID 068032 PRELIMINARY; PRT: 212 AA.
AC 068032;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CHIGINE-2'86;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85296; G736139; -
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24308 MW; 69D87B53 CRC32.

Query Match 65.6%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 4,30e-225;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 88
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 148
QY 155 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 214

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
QY 215 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 274

Db 209 ESOC 212
QY 275 ESOC 278

QY 155 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 214
Db 149 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
QY 215 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 274

Db 209 ESOC 212
QY 275 ESOC 278

RESULT 14
ID 068010 PRELIMINARY; PRT: 212 AA.
AC 068010;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100; -
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24292 MW; 04A4D12D CRC32.

Query Match 65.5%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 6,64e-225;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 88
QY 95 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHHTALROAI 154

Db 89 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 148
QY 155 LCMGELMTLATWGVNLEDASRDLYVSYVNTNMGKLFROLMFHISCLTFGRVTEYL 214

Db 149 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
QY 215 VSEGVWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 274

Db 209 ESOC 212
QY 275 ESOC 278

RESULT 15
ID 068070 PRELIMINARY; PRT: 212 AA.
AC 068070;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GIORDO'84;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210; -
DR PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24376 MW; 80F52D0F CRC32.

Query Match 65.4%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1,58e-224;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

US3890-10-38183
MDIDPYKEFGATVEELSLFSPDFFPSVDLDTASALYREALSPHHTALROAILCWEELMTLAT
WGVNLEDPASRDVVSYNLLQMDFGPEHLVDLQSLMDIDPYKEFGATVEELSLFSPDFFPSVR
DLDTASALYREALSPHHTALROAILCWEELMTLATWGVNLEDPASRDVVSYNLLQMDLKER
OLMFIHISCLTFEGRETVIEYLVFCWVIRTPPAYRPNAPILSTLPETTIVRRGRSPRRRTPPRRRS
OSPRRRRSQSRSSQ1

 M I S E R I E

 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch:dp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:44:26 1999; MasPar time 13.43 Seconds

Tabular output not generated. 467.254 Million cell updates/sec

Title: >US3890-10-38183
 Description: (1-295) from us3890-10-38183.pep
 Perfect Score: 2205
 Sequence: 1 MDIDYKEFGATVELSLFP.....RRRSQSPRRRSQSHSQC 295

Scoring table: PAM 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.363; Variance 173.459; scale 0.192

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	1375	62.4	194	32	W50242	Hepatitis B virus pre
2	1375	62.4	212	32	W50250	Hepatitis B virus pre
3	1375	62.4	346	5	R27473	S12/core protein.
4	1363	61.8	184	1	P80959	Hepatitis B virus sub
5	1358	61.6	183	32	W50251	Hepatitis B virus p21
6	1359	61.6	193	32	W50241	Hepatitis B virus pre
7	1353	61.4	183	20	W09044	Hepatitis B virus cor
8	1349	61.2	183	1	R05635	Hepatitis B antigen.
9	1349	61.2	183	5	P00041	Sequence of core anti
10	1344	61.0	397	20	W09048	Plasmodium PHV DN AA en
11	1338	60.7	184	5	P00004	Sequence of core anti
12	1336	60.6	196	8	R40806	Hepatitis B core / PV
13	1337	60.6	208	8	R40808	Hepatitis B core / PV
14	1334	60.5	183	4	P40311	Hepatitis B virus prot
15	1334	60.5	183	8	R40805	Hepatitis B core prot
16	1334	60.5	183	13	R68868	Hepatitis B virus pol

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	DB ID	Description	Pred. No.
1	W50242	60.5	183	12	R62869	Hepatitis B virus cor	2.91e-96
2	W50242	60.5	198	8	R40807	Hepatitis B core / TL	3.32e-96
3	W50242	60.2	293	10	R5286	Presi full length cor	9.02e-96
4	W50242	60.1	183	13	R68866	Hepatitis B virus pol	1.59e-95
5	W50242	60.1	214	1	P90702	Deduced amino acid se	1.59e-95
6	W50242	60.1	289	20	W09049	Plasmodium PHV DN BB en	1.59e-95
7	W50242	60.0	185	6	R30861	Hepatitis B core anti	1.92e-95
8	W50242	60.0	185	10	R35284	Deduced sequence of f	1.92e-95
9	W50242	60.0	185	6	R30780	Native HBcAg protein.	1.92e-95
10	W50242	60.0	185	6	R31025	Native HBcAg protein.	1.92e-95
11	W50242	59.9	183	13	R68869	Hepatitis B virus pol	4.07e-95
12	W50242	59.9	185	26	P30061	Core antigen of HBV.	3.37e-95
13	W50242	59.7	185	6	R30781	Altered HBcAg protein	7.16e-95
14	W50242	59.7	185	6	R31026	Altered HBcAg protein	7.16e-95
15	W50242	59.6	183	13	R68867	Hepatitis B virus pol	1.26e-94
16	W50242	59.5	183	13	R68870	Hepatitis B virus pol	1.52e-94
17	W50242	59.5	185	6	R30784	Altered HBcAg protein	2.21e-94
18	W50242	59.5	185	6	R31029	Altered HBcAg protein	2.21e-94
19	W50242	59.5	185	6	R30865	Modified hepatitis B	2.21e-94
20	W50242	59.5	185	6	R30865	Modified hepatitis B	2.21e-94
21	W50242	59.4	185	6	R31028	Hepatitis B core anti	3.23e-94
22	W50242	59.3	185	6	R30783	Altered HBcAg protein	3.89e-94
23	W50242	59.3	185	6	R30864	Modified hepatitis B	3.89e-94
24	W50242	59.1	214	1	P80961	HBV core antigen enco	9.98e-94
25	W50242	59.0	203	1	P82872	HBcAg/beta-gal fusion	1.20e-93
26	W50242	58.4	198	32	W50252	Hepatitis B virus pre	2.03e-92
27	W50242	57.3	183	17	R68878	Hepatitis B virus E	1.85e-90
28	W50242	57.1	183	17	R68883	Hepatitis B virus E	3.99e-90
29	W50242	57.1	183	17	R68884	Hepatitis B virus E	4.74e-90

RESULT 1
 ID W50242 standard; Protein; 194 AA.

AC W50242; 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p22 polypeptide Met-p22.

KW Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
 KW hepatocyte; liver; Met-p22.

OS Hepatitis B virus.
 OS Synthetic.

FT Key
 FT Protein

FT Location/Qualifiers
 FT 2..194 /Label= p22

PD W0909649-A1.
 PD 12-MAR-1998.

PF 03-SEP-1997; U15500.
 PF 03-SEP-1996; US-025370.

PA (GENE) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;

DR WPI: 98-19325/17.

PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus

PS Claim 11; Page 40; 60pp; English.

CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein with an added N-terminal Met residue; p22 is produced by
 CC elimination of the 19-amino acid leader peptide from the 25 kDa

CC full-length HBV precore protein (see W50250). Evidence is provided
 CC that HBV replication is inhibited in the presence of high levels of
 CC HBV precore or precore-related proteins. These proteins can be

CC incorporated into HBV nucleocapsids along with the p1 core protein
 CC (see W50251), which is the usual nucleocapsid component, and
 CC thereby render the nucleocapsids deficient in encapsidating HBV

CC pregenomic RNA. Thus, over-expression of the precore proteins, or
 CC certain variants of them, leads to transdominant inhibition of HBV

CC replication. Suitable inhibitory proteins include p25 (see W50250),
 CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het

CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory

CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously

CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 194 AA:

Query Match 62.4%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 1.29e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmdidpkykefgatvelisfipdsffsvrdlildtasalyrealsephscphntalrgai 70
 :|||||
 QY 112 SMDIDPKKEGATVELLSFIPSPDFPSVRDLDTASALYREALSEPHSCPHNTALRQAI 171
 :|||||
 Db 71 lwcgelmclatwgvnledpdsrdlvsyvnltmgikfqiqlwfhiscitfgretvleyl 130
 :|||||
 QY 172 LCGELMTLATWGVNLEDPASRDLVSVYNTMGLKFRQLMFHISCITFGRETVEYL 231
 :|||||
 Db 131 vsfgwvltreppayrppnnapilslpetvrrrgsrprrrrrsgsprrrrrsgsr 190
 :|||||
 QY 232 VSGFWIRTPPAYRPPNAPILSLPETVRRRGSRPRRRTPSPRRRSQSPRRRSQSR 291
 :|||||
 Db 191 esgc 194
 :|||||
 QY 292 ESQC 295

RESULT 2
 ID W50250 standard; Protein: 212 AA.

AC W50250:
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p25 polypeptide.
 KW Hepatitis B virus replication; inhibitor; HBV; nucleocapsid; gene therapy;
 KW hepatocyte; liver; p25 protein.
 OS Hepatitis B virus.
 PN MO9809649-A1.
 PD 12-MAR-1998.
 PE 03-SEP-1997; 015500.
 PR 03-SEP-1996; US-025370.
 PA (GENE) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni P, Wands JR;
 DR WPI: 98-193325/17
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 15: Page 35; 60pp; English.
 CS This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SQ Sequence 212 AA:

Query Match 62.4%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 1.29e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmdidpkykefgatvelisfipdsffsvrdlildtasalyrealsephscphntalrgai 88
 :|||||
 QY 112 SMDIDPKKEGATVELLSFIPSPDFPSVRDLDTASALYREALSEPHSCPHNTALRQAI 171
 :|||||

Db 89 lwcgelmclatwgvnledpdsrdlvsyvnltmgikfqiqlwfhiscitfgretvleyl 148
 :|||||
 QY 172 LCGELMTLATWGVNLEDPASRDLVSVYNTMGLKFRQLMFHISCITFGRETVEYL 231
 :|||||

Db 149 vsfgwvltreppayrppnnapilslpetvrrrgsrprrrrrsgsprrrrrsgsr 208
 :|||||
 QY 232 VSGFWIRTPPAYRPPNAPILSLPETVRRRGSRPRRRTPSPRRRSQSPRRRSQSR 291
 :|||||

Db 209 esgc 212
 :|||||
 QY 292 ESQC 295

RESULT 3
 ID R27473 standard; Protein: 346 AA.

AC R27473:
 DT 24-FEB-1993 (first entry)
 DE S12/core protein.
 KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KW vaccinia virus; I3L; promoter; NVVAC; recombinant; HBV L;
 KW large pre-S antigen; IspAg; fusion protein; pre-S region; S12/core;
 KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KW deletion loci; recipient loci.
 OS Synthetic.
 FH Key
 FT Location/Qualifiers
 FT region 1..108
 FT /label S1
 FT region 109..163
 FT /label S2
 FT region 164..346
 FT /label Core
 PN WO9215672-A.
 PD 17-SEP-1992.
 PE 09-MAR-1992; U01906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRO-) VIROGENETICS CORP.
 PI COX MI, De Taisne C, Francis J, Gettig RR, Johnson GP,
 PI Limbach KJ, Norton EK, Paoletti E, Perkins ME, Plancus SE,
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI: 92-331718/40.
 PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure: Fig 13; 45pp; English.

CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the I3L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NVVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, IspAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NVVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NVVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC 035501-864.
 SQ Sequence 346 AA:

Query Match 62.4%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 1.29e-99;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 mndidpkykefgatvelisfipdsffsvrdlildtasalyrealsephscphntalrgai 222
 :|||||

Page 3

DT	28-SEP-1998	(first entry)	
DE	Hepatitis B virus p21 core protein.		
DD	Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;		
KW	hepatocyte; liver; p21; core protein.		
KW	Hepatitis B virus.		
OS	key		
FH	Location/Qualifiers		
FT	72..88		
FT	Region		
PN	MO9809649-A1.		
PN	12-MAR-1998.		
PD	03-SEP-1997.	U15500.	
PF	03-SEP-1996.	US-025370.	
PR	(GEHO) GEN HOSPITAL CORP.		
PA	Melegari M, Scagliioni PP, Wands JR;		
IC1			

	PT	nucleocapsid subunit(s) into a viral nucleocapsid - useful for
	PT	Inhibition of viral replication, especially hepatitis B virus
	PS	Disclosure, Page 41-42; 60pp; English.
	CC	This is the 21 kDa (p21) protein of hepatitis B virus (HBV) that
	CC	assembled into a 180 Ksa subunit nucleocapsid structure that
	CC	Promotes viral replication. Evidence is provided that HBV
	CC	replication is inhibited in the presence of high levels of HBV
	CC	precure or precure-related proteins. These proteins can be
	CC	incorporated into HBV nucleocapsids along with the p21 core protein
	CC	and thereby render the nucleocapsid deficient in encapsidating HBV
	CC	pregenomic RNA. Thus, over-expression of the precure proteins or
	CC	certain variants of them, leads to transdominant inhibition of HBV
	CC	replication. Suitable inhibitory proteins include p25 (see
	CC	W50230), p22 (see W50241), Met-p22 (see W50242), p18 (see
	CC	Met-p18 (see W50237) and Met-p18-Het (see W50238). These proteins
	CC	can be produced by recombinant methods using claimed expression
	CC	vectors and host cells, and can be provided exogenously to target
	CC	cells for use in inhibiting HBV replication. Alternatively, a
	CC	nucleic acid construct that directs overexpression of an inhibitory
	CC	protein in target cells is used for the gene therapy of HBV
	CC	infection.
SQ	Sequence	183 AA:
	Query Match	61.6%; Score 1358; DB 32; Length 183;
	Best Local Similarity	98.9%; Pred. No. 3,17e-98;
	Matches 181; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
Dd	1	mididyrefatavllsfpsdffprvrdltdtsalyprealsephsphtlrgall 60
OY	113	MDIDPRKEFGATVEELFFLSDFSPFRKDLCDIRASALRELBSEPHCSPHITAKQAIL 172
	61	cwgelmlacwgygnledpasrdilvsvydtmnglkfrqllwfmscltfgretylelv 120

QY	173	CWGLMIMLAIVMGVNLDDPSRDLVSVYNTNMGLKFKFQGLMLFHISCLTFGEETVIEYLV	232
Db	121	sfgvwlittpayyirppnaqlstlptectvvirrgtgsprirtcpsprirrrzsqaprrirzsgare	180
QY	233	sfgvwmirtppayirppnaqlstlptectvvirrrrgsrprrrrrpsprrrrsqsprrrrsosre	292
Db	181	sgc 183	
QY	293	sqc 295	
DT	28-SEP-1998	(first entry)	
KW	Hepatitis B virus precore p22 polypeptide.		
KW	viral replication; inhibitor; HBV; nucleocapsid; gene therapy;		
KW	hepatocyte; liver; p22 protein.		
OS	Hepatitis B virus.		
EH	key	location/Qualifiers	
FT	Region	82..98	
PN	WO9809649-A1.	/note="Immunodominant region"	
PD	12-MAR-1998.		

PF 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (GHEO) GEN HOSPITAL CORP.
 PI Melegari M. Scaglioni PP, Wands JR:
 DR WPI; 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 9: Page 34-35; 60pp: English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein that is produced by elimination of the 19-amino acid leader
 CC peptide from the 25 kDa full-length HBV precore protein (see
 CC W50250). Evidence is provided that HBV replication is inhibited in
 CC the presence of high levels of HBV precore or precore-related
 CC proteins. These proteins can be incorporated into HBV nucleocapsids
 CC along with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25 (see W50250), p22, Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
 CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously
 CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 193 AA;

Query Match 61.6%; Score 1359; DB 32; Length 193;
 Best Local Similarity 98.4%; Pred. No. 2.63e-98;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 10 gmdipkyefgatvellstfspdffsvrdldltasalyrealespeshcphntalrqal 69
 :|||||
 QY 112 SMDIDYKEFGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHNTALRQAI 171
 Db 70 lcgwelmrlatwgvnledpasrdlvsvyrdtmglkfrqlwfhiscldfgetvleyl 129
 :|||||
 QY 172 LCGWELMTLATWGVNLEDPASRDLVSVYNTNMGKFRQLWFMHISCLFGETVLEYL 231
 Db 130 vsfgwvrltppayrppnapllstlpettvrrrgsrprrrrrsgsrrrrsgsr 189
 :|||||
 QY 232 VSGWVIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 291
 Db 190 esgc 193
 :||||
 QY 292 ESQC 295

RESULT 7
 ID W09044 standard; Protein: 183 AA.
 AC W09044;
 DT 11-APR-1997 (first entry)
 DE Hepatitis B virus core protein.
 DE Hepadnavirus; HBV; woodchuck hepatitis virus; hepatitis delta virus;
 KW ground squirrel hepatitis B virus; duck hepatitis B virus;
 KW core protein; replication; antiviral; gene therapy.
 OS Hepatitis B virus.
 FH Hepatitis B virus.
 FT misc-difference 71..180
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 71-180 in mutant
 FT polypeptides of the invention (Claim 23)"
 FT 81..180
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 81 and 180 in mutant
 FT polypeptides of the invention (Claim 5)"
 FT 171..180
 FT /note= "C-terminus of core protein is at any amino

FT acid position between 171 and 180 in mutant
 FT polypeptides of the invention (Claim 6)"
 FT 174..180
 FT /note= "C-terminus of core protein is at any amino
 FT acid position between 174 and 180 in mutant
 FT polypeptides of the invention (Claim 26)"
 FT 178
 FT /note= "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT 178
 FT /note= "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT 178
 FT /note= "C-terminus of core protein is at amino acid
 FT position 178 in mutant polypeptides of the
 FT invention (Claim 7)"
 FT 172..183
 FT /note= "amino acid residues 172-183, pref. 174-180,
 FT are deleted from the core protein in mutant
 FT polypeptides of the invention"

PN W09700698-A1.
 PD 09-JAN-1997.
 PE 20-JUN-1996; U10602.
 PF 20-JUN-1995; US-017814.
 PA (GHEO) GEN HOSPITAL CORP.
 PI Melegari M. Scaglioni PP, Wands JR:
 DR WPI; 97-087176/08.
 PI N-PSDB; T49594.

PT New method for inhibiting the replication of hepadnaviruses -
 PT comprises introducing a mutant polypeptide with a mutated core
 PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
 PS B
 PS Claim 5; Page 55-56; 83pp: English.
 CC Alteration of the C-terminus of the core protein (W09044) of
 CC hepatitis B virus (HBV) or other hepadnavirus creates a mutant
 CC polypeptide capable of reducing replication of the wild-type virus
 CC by a dominant negative mechanism. The inhibitory effect is
 CC species-specific and is achieved by deletion of a few C-terminal
 CC amino acids from the core protein, and/or by joining the core
 CC protein to a hepadnavirus surface protein (see also W09045),
 CC creating a core-surface fusion. Such mutant polypeptides (see
 CC also W09046-50) can be used to treat hepadnavirus, e.g. HBV,
 CC infection.
 SQ Sequence 183 AA;

Query Match 61.4%; Score 1353; DB 20; Length 183;
 Best Local Similarity 98.9%; Pred. No. 8.14e-98;
 Matches 181; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 mididpykefgatvellstfspdffsvrdldltasalyrealespeshcphntalrqal 60
 :|||||
 QY 113 MDIDPYKEFGATVELLSFSPDFPSVRDLDTASALYREALESPHCSPHNTALRQAIL 172
 Db 61 cwgelmrlatwgvnledpasrdlvsvyrdtmglkfrqlwfhiscldfgetvleyl 120
 :|||||
 QY 173 CWGELMTLATWGVNLEDPASRDLVSVYNTNMGKFRQLWFMHISCLFGETVLEYL 232
 Db 121 sfvgwvrltppayrppnapllstlpettvrrrgsrprrrrrsgsrrrrsgsre 180
 :|||||
 QY 233 SFGWVIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 292
 Db 181 sqc 183
 :|||
 QY 293 SQC 295

RESULT 8
 ID R05635 standard; protein: 183 AA.
 AC R05635;
 DT 30-OCT-1990 (first entry)
 DE Hepatitis B antigen.
 DE Hepatitis B virus; vaccine; HBV; ds.
 KW Hepatitis B virus; vaccine; HBV; ds.
 OS Synthetic.
 PN EP-374869-A.
 PD 27-JUN-1990.

PF 1-JAN-1989; 123526.
PR 22-DEC-1978; GB-049907.
PR 27-DEC-1978; GB-050039.
PR 01-NOV-1979; GB-037910.
PA (BIOJ) Biogen Inc.
PI Murray K. Schaller HE;
DR WPI: 90-195067/26.
N-PSDB: 004799.
PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
PS used in detection of infection and in vaccine prodn.
CC Disclosure: 4pp: English.
CC Polypeptide fragments encoded by the DNA sequence are antigenic for
CC HBV and may be used as a vaccine or in detection. Peptides may be
CC cultured in a suitable bacterial host such as E.coli.
CC Fragments of the sequence are also claimed as being antigenically
CC useful.
SQ Sequence 183 AA:

Query Match 61.2%; Score 1349; DB 1; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.73e-97;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdipykefgatvelisfipdsffpsvrdldtaaalrydalespeshphthalrgail 60
|||
QY 113 MDIDPYKEFGATVELISFIPDSFFPSVARDLDTASALYREALSEPHSPHTALRQAIL 172
|||
Db 61 cwgdlmtlatwgvnlgedpasrdlvsvyvnvmgkrfqlwfhlscltfgretvleylv 120
|||
QY 173 CWGELMTLATWGVNLGEDPASRDLVSVYVNTNMGKRFQLWFMHISCLTFGRETVIEYLV 232
|||
Db 121 sfgywlrtppayrppnapilslpetvtvrrrgsprrrtsprrrrsgsrrrrsgsre 180
|||
QY 233 SFGWIRTPPAYRPPNAPILSLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSRE 292
|||
Db 181 sqc 183
|||
QY 293 SQC 295

RESULT 9
ID P00041 standard; Protein; 183 AA.

AC P00041; 1992 (first entry)
DT 14-OCT-1992
DE Sequence of core antigen.
KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
OS Hepatitis B virus.
PN EP-13828-A.
PD 06-AUG-1980.
PF 21-DEC-1979; 303017.
PR 22-DEC-1978; GB-049907.
PR 27-DEC-1978; GB-050039.
PR 01-NOV-1979; GB-037910.
PA (BIOJ) Biogen Inc.
PI Murray K. Schaller HE;
DR WPI: 80-57268C/33.
N-PSDB: N00003.
PT Recombinant DNA coding for polypeptide - have specificity of
PT hepatitis B viral antigens in detection or antibody stimulation
PS Example: Figs 3-4; 43pp: English.
CC Human serum from a single HBsAg positive, HBsAg positive donor
CC (serotype adym) was used to prep. a DNA-contg. pellet which was
CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
CC The labelled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prepd. by the processes are deposited at the
CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg:
CC HBV-Kpn I dg: Tetr Amps HBV+.
SQ Sequence 183 AA:

Query Match 61.2%; Score 1349; DB 5; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.73e-97;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdipykefgatvelisfipdsffpsvrdldtaaalrydalespeshphthalrgail 60
|||
QY 113 MDIDPYKEFGATVELISFIPDSFFPSVARDLDTASALYREALSEPHSPHTALRQAIL 172
|||
Db 61 cwgdlmtlatwgvnlgedpasrdlvsvyvnvmgkrfqlwfhlscltfgretvleylv 120
|||
QY 173 CWGELMTLATWGVNLGEDPASRDLVSVYVNTNMGKRFQLWFMHISCLTFGRETVIEYLV 232
|||
Db 121 sfgywlrtppayrppnapilslpetvtvrrrgsprrrtsprrrrsgsrrrrsgsre 180
|||
QY 233 SFGWIRTPPAYRPPNAPILSLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSRE 292
|||
Db 181 sqc 183
|||
QY 293 SQC 295

RESULT 10
ID W09048 standard; Protein; 397 AA.

AC W09048;
DT 11-APR-1997 (first entry)
DE Plasmid pHBV DN AA encoded HBV dominant negative polypeptide.
KW Hepadnavirus; woodchuck hepatitis virus; hepatitis delta virus;
KW HBV; ground squirrel hepatitis B virus; duck hepatitis B virus;
KW core protein; replication; antiviral; gene therapy; pHBV DN AA.
OS Hepatitis B virus.
FH Key
FT region 1.179
FT location/Qualifiers
FT 1.179
FT /label- HBV_core
FT /note- "positions 1-179 correspond to amino acids
FT 1-179 of HBV core protein"
FT 180..397
FT /note- "positions 180-397 correspond to amino
FT acids 9-226 of HBV surface protein"

PN W09700698-A1.
PD 09-JAN-1997.
PF 20-JUN-1996; U10602.
PR 20-JUN-1995; US-017814.
PA (GENO.) GEN HOSPITAL CORP.
PI Melegari M. Scaglioni PP, Wands JR.
DR WPI: 97-087176/08.
N-PSDB: T49598.
PT New method for inhibiting the replication of hepadnaviruses -
PT comprises introducing a mutant polypeptide with a mutated core
PT protein or corresponding nucleic acid, for treating, e.g. hepatitis
PT B.
PS Disclosure: Page 46-48; 83pp: English.
CC Plasmid pHBV DN AA carries an insert (T49598) coding for a protein
CC (W09048) comprising a hepatitis B virus (HBV) core protein (see
CC also W09044) fused in-frame at amino acid 179 with the HBV surface
CC protein (see also W09045) at amino acid 8. Plasmid pHBV DN BB
CC (T49599) expresses the HBV core fused at amino acid 175 to the
CC surface protein at amino acid 112 (W09049). pHBV DN AA was at
CC least as potent an inhibitor of HBV replication as construct
CC pHBV DN (T49597); pHBV DN BB was less inhibitory than pHBV DN.
CC Vectors expressing hepadnavirus dominant negative core mutants can
CC be utilised in the gene therapy of viral infections.
SQ Sequence 397 AA:

Query Match 61.0%; Score 1344; DB 20; Length 397;
Best Local Similarity 100.0%; Pred. No. 4.43e-97;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdipykefgatvelisfipdsffpsvrdldtaaalrydalespeshphthalrgail 60
|||
QY 113 MDIDPYKEFGATVELISFIPDSFFPSVARDLDTASALYREALSEPHSPHTALRQAIL 172
|||
Db 61 cwgdlmtlatwgvnlgedpasrdlvsvyvnvmgkrfqlwfhlscltfgretvleylv 120
|||
QY 173 CWGELMTLATWGVNLGEDPASRDLVSVYVNTNMGKRFQLWFMHISCLTFGRETVIEYLV 232
|||
Db 121 sfgywlrtppayrppnapilslpetvtvrrrgsprrrtsprrrrsgsrrrrsgsre 179
|||

QY 233 SFGWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291

RESULT 11
ID P00004 standard: Protein; 184 AA.

AC P00004; 14-OCT-1992 (first entry)
DE Sequence of core antigen.
KW Hepatitis B virus; antigen; antibody; diagnosis; vaccine.
OS Hepatitis B virus.

PN EP-13828-A.
PD 06-AUG-1980.
PF 21-DEC-1979; 303017.
PR 22-DEC-1978; GB-049907.
PR 27-DEC-1978; GB-050039.
PR 01-NOV-1979; GB-037910.
PA (BIOT) BIOGEN NV.
PI Murray K. Schaller HE;
DR WPI: 80-57268C/33.
N-PSDB: N00002.
PT Recombinant DNA coding for polypeptide - have specificity of
PT hepatitis B viral antigens in detection or antibody stimulation
PS Claim 13; Page 40; 43pp; English.
CC Human serum from a single HBsAg positive, HBeAg positive donor
CC (serotype adym) was used to prep. a DNA-contg. pellet which was
CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
CC The labelled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prepd. by the processes are deposited at the
CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg:
CC HBV-Kpn I dc: Tetr Amps HBV+.
SQ Sequence 184 AA;

Query Match 60.7%; Score 1338; DB 5; Length 184;

Best Local Similarity 96.2%; Pred. No. 1.37e-96;
Matches 177; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

DB 1 mdidpykefgasvellsfipdsffpsvordlldtaaalysalesepheosphhtalrqa 60
QY 113 MDIDPYKEFGATVELLSFLPSDFPSPV-RDLDYKASALYREALSEPHCSPHHTALRQA 171
DB 61 lcgwlmmlatwvgnledpasrdlvsvyvmnglklrqlwfhisccltfgretvleyl 120
QY 172 LCGWELMTLATWVGNLEDPASRDLVSVYVNTMNGLKFROLWFHISCCLTFGRETVEYL 231
DB 121 vsfgvwtppayrppnapilstlpettvrrgrspprrrtsprrrsosprrrrsosr 180
QY 232 VSFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
DB 181 esgc 184
QY 292 ESQC 295

RESULT 12
ID R40806 standard: Protein; 196 AA.

AC R40806; 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 fusion.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FH Key Location/Qualifiers
FT region 7..13
FT /label= PV-1
FT protein 14..196
FT /label= HBC

PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (N1SW) NISSHIN OIL MILLS LTD.

DR WPI: 93-277479/35.

DR N-PSDB: Q47736.

PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes

PS Disclosure; Fig 7; 12pp; Japanese.

CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.

SQ Sequence 196 AA;

Query Match 60.6%; Score 1336; DB 8; Length 196;

Best Local Similarity 95.1%; Pred. No. 2.00e-96;
Matches 175; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 13 amdidpykefgasvellsfipdsffpsvordlldtaaalysalesepheosphhtalrqa 72
QY 112 SMDIDPYKEFGATVELLSFLPSDFPSPV-RDLDYKASALYREALSEPHCSPHHTALRQA 171
DB 73 lcgwlmmlatwvgnledpasrdlvsvyvmnglklrqlwfhisccltfgretvleyl 132
QY 172 LCGWELMTLATWVGNLEDPASRDLVSVYVNTMNGLKFROLWFHISCCLTFGRETVEYL 231
DB 133 vsfgvwtppayrppnapilstlpettvrrgrspprrrtsprrrsosprrrrsosr 192
QY 232 VSFGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
DB 193 esgc 196
QY 292 ESQC 295

RESULT 13
ID R40808 standard: Protein; 208 AA.

AC R40808; 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 / IL-1 fusion.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FH Key Location/Qualifiers
FT region 8..19
FT /label= PV-1
FT region 20..25
FT /label= IL-1
FT protein 26..208
FT /label= HBC

PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (N1SW) NISSHIN OIL MILLS LTD.
DR WPI: 93-277479/35.
N-PSDB: Q47738.
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes
PS Disclosure; Fig 9; 12pp; Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.

SQ Sequence 208 AA;

Query Match 60.6%; Score 1337; DB 8; Length 208;

Best Local Similarity 94.6%; Pred. No. 1.66e-96;
Matches 175; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

DB 24 lmgdidpykefgasvellsfipdsffpsvordlldtaaalysalesepheosphhtalrqa 83
QY 111 LMGDIDPYKEFGATVELLSFLPSDFPSPV-RDLDYKASALYREALSEPHCSPHHTALRQA 170
DB 84 lcgwlmmlatwvgnledpasrdlvsvyvmnglklrqlwfhisccltfgretvleyl 143

```

OY 171 ILGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMHISCLTFGRETVLEY 230
DB 144 LVSGVWIRTPPAYRPPNAPILSTLPETVVRIRGRSPRRTPSPRRRRSQRSS 203
OY 231 LVSGVWIRTPPAYRPPNAPILSTLPETVVRIRGRSPRRTPSPRRRRSQRSS 290
DB 204 resqc 208
OY 291 RESQC 295

```

```

RESULT 14
ID P40311 standard; Protein; 183 AA.
AC P40311:
DT 16-AUG-1992 (first entry)
DE Hepatitis virus core antigen.
KW HbcAg; vaccine; diagnosis; HBV infection.
OS Hepatitis B virus.
PN J59074985-A.
PD 27-APR-1984.
PF 19-OCT-1982; 183432.
PR 19-OCT-1982; JP-183432.
PA (TAKE ) TAKEDA CHEMICAL IND KK.
DR WPI: 84-143231/23.
PT DNA used in prevention of infections by hepatitis virus B -
PT comprises structural gene of hepatitis virus adr B surface antigen
PS Disclosure: Fig 3; 13pp; Japanese.
CC The sequence is that of hepatitis virus core antigen (HbcAg). It
CC can be used as a vaccine for the prevention of infections by
CC hepatitis B virus (HBV) and also in the diagnosis of early stages
CC of HBV infection. See also P40310.
SQ Sequence 183 AA:

```

```

Query Match 60.5%; Score 1334; DB 4; Length 183;
Best Local Similarity 95.6%; Pred. No. 2,91e-96;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdidpykefgasvellsfipdsdfpsirdltdtasalyrealespeshcphhtalrga1 60
OY 113 MDIDPYKEFGATVELLSFSDFFPSVRDLDTASALYREALSPESHCPHHTALRQA1L 172
DB 61 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMHISCLTFGRETVLEY1V 120
OY 173 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMHISCLTFGRETVLEY1V 232
DB 121 sfgywlrtppayrppnapilstlpetvvrirgrspprrtpsprrrrsqrssre 180
OY 233 SFGVWIRTPPAYRPPNAPILSTLPETVVRIRGRSPRRTPSPRRRRSQRSSRE 292
DB 181 sqc 183
OY 293 SQC 295

```

```

RESULT 15
ID R40805 standard; Protein; 183 AA.
AC R40805:
DT 16-FEB-1994 (first entry)
DE Hepatitis B core protein.
KW Hepatitis B; core; HBC; PV-1; IL-1 beta; Interleukin; plasmid.
OS Hepatitis B virus.
PN J05192170-A.
PD 03-AUG-1993.
PF 24-SEP-1991; 243800.
PR 24-SEP-1991; JP-243800.
PA (NISM ) NISSHIN OIL MILLS LTD.
DR WPI: 93-277479/35.
DR N-PSDB: 047735.
PT Recombinant plasmid for high immunogenicity virus - contains
PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia
PT virus and exotic genes

```

```

PS Disclosure: Fig 6; 12pp; Japanese.
CC The hepatitis B core gene is recombinant with PV-1 DNA
CC and IL-1 beta to form a plasmid (Q47738) which is then used to produce
CC a vaccine. NB: Sequence is difficult to read in the original
CC specification.
SQ Sequence 183 AA:

```

```

Query Match 60.5%; Score 1334; DB 8; Length 183;
Best Local Similarity 95.6%; Pred. No. 2,91e-96;
Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

DB 1 mdidpykefgasvellsfipdsdfpsirdltdtasalyrealespeshcphhtalrga1 60
OY 113 MDIDPYKEFGATVELLSFSDFFPSVRDLDTASALYREALSPESHCPHHTALRQA1L 172
DB 61 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMHISCLTFGRETVLEY1V 120
OY 173 CWGELMTLATWGVNLEDPASRDVSVYNTNMGKLFROLMHISCLTFGRETVLEY1V 232
DB 121 sfgywlrtppayrppnapilstlpetvvrirgrspprrtpsprrrrsqrssre 180
OY 233 SFGVWIRTPPAYRPPNAPILSTLPETVVRIRGRSPRRTPSPRRRRSQRSSRE 292
DB 181 sqc 183
OY 293 SQC 295

```

```

Search completed: Thu Dec 16 13:47:38 1999
Job time : 192 secs.

```


Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

MPsrch_pp  protein - protein database search, using Smith-Waterman algorithm
Run on:      Thu Dec 16 13:47:54 1999;  MasPar time 14.55 Seconds
Tabular output not generated.      812.638 Million cell updates/sec

```

```

Title:                >US3890-10-38183
Description:           (1-295) from us3890-10-38183.pep
Perfect Score:        2205
Sequence:             1 MDIDPYEFGATYELLFLD.....RRRRSQSPRRRRSQSRSSQC 295

```

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

```
Listing first 45 summaries
plr60
1:plr1 2:plr2 3:plr3 4:plr4
```

Statistics: Mean 47.101; Variance 117.991; scale 0.3999

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1375	62.4	212	1	NKVLAH	e antigen precursor /	8.67e-2011
2	1373	62.3	183	2	S53207	core antigen - hepati	1.88e-2000
3	1374	62.3	212	2	S53211	e antigen precursor /	1.28e-2000
4	1374	62.3	212	2	S32204	e antigen precursor /	1.28e-2000
5	1373	62.3	212	2	S07050	e antigen precursor /	1.88e-2000
6	1370	62.1	183	2	S53200	core antigen - hepati	5.98e-2000
7	1370	62.1	212	2	S53216	e antigen precursor /	5.98e-2000
8	1370	62.1	212	2	S53212	e antigen precursor /	5.98e-2000
9	1369	62.1	212	2	S33163	e antigen precursor /	8.80e-2000
10	1369	62.1	212	2	S20746	e antigen precursor /	8.80e-2000
11	1369	62.1	212	2	S53225	e antigen precursor /	8.80e-2000
12	1368	62.0	212	2	S53198	e antigen precursor /	1.30e-1999
13	1366	62.0	212	2	S53223	e antigen precursor /	2.80e-1999
14	1366	62.0	212	2	S53274	e antigen precursor /	2.80e-1999
15	1364	61.9	183	2	S53169	core antigen - hepati	6.07e-1999
16	1363	61.8	183	2	S53247	e antigen - hepati	8.93e-1999
17	1363	61.8	212	2	S53202	e antigen precursor /	8.93e-1999
18	1363	61.8	212	2	S53159	e antigen precursor /	8.93e-1999
19	1363	61.8	212	2	S53200	e antigen precursor /	8.93e-1999
20	1362	61.8	212	2	S53251	e antigen precursor /	1.31e-1998
21	1360	61.7	212	2	S53204	e antigen precursor /	2.84e-1988
22	1360	61.7	212	2	S53227	e antigen precursor /	2.84e-1988
23	1358	61.6	211	1	NKVLAH	e antigen precursor /	6.16e-1988

45	1337	60.6	212	1	NKVIL2	e antigen precursor /	4.18e-198
44	1338	60.7	212	2	SS3236	e antigen precursor /	9.06e-198
43	1339	60.7	212	2	SS3236	e antigen precursor /	1.96e-197
42	1339	60.7	183	2	SS3137	e antigen precursor /	1.96e-197
41	1340	60.8	212	2	SS3237	e antigen precursor /	9.18e-197
40	1341	60.8	183	2	SS3257	e antigen precursor /	1.96e-197
39	1341	60.8	183	2	SS3159	e antigen precursor /	1.96e-197
38	1341	60.8	183	2	SS3257	e antigen precursor /	1.96e-197
37	1342	60.9	183	2	SS3210	e antigen precursor /	1.96e-197
36	1343	60.9	183	2	SS3140	e antigen precursor /	1.96e-197
35	1344	61.0	212	2	SS3229	e antigen precursor /	1.96e-197
34	1345	61.0	183	2	SS3152	e antigen precursor /	1.96e-197
33	1348	61.1	183	2	SS3260	e antigen precursor /	1.96e-197
32	1348	61.1	183	2	SS3222	e antigen precursor /	1.96e-197
31	1349	61.2	183	2	SS3181	e antigen precursor /	1.96e-197
30	1349	61.2	183	1	NKVIL2	e antigen precursor /	1.96e-197
29	1351	61.3	212	2	SS3240	e antigen precursor /	1.96e-197
28	1351	61.3	183	2	SS3139	e antigen precursor /	1.96e-197
27	1355	61.5	212	1	NKVIL2	e antigen precursor /	1.96e-197
26	1355	61.5	212	1	NKVIL2	e antigen precursor /	1.96e-197
25	1357	61.5	212	2	SS3231	e antigen precursor /	1.96e-198
24	1359	61.6	212	2	SS3233	e antigen precursor /	1.96e-198

ALIGNMENTS

```

RESULT      1
ENTRY
TITLE      NKVLAH      #type complete
            e antigen precursor / core antigen - hepatitis B virus
            (subtype ayw4, isolate hb321 and others)
ALTERNATE_NAMES
CONTAINS    HBe antigen precursor / Hbc antigen; pre-C/C antigen
            core antigen; e antigen
ORGANISM    #formal name hepatitis B virus; HBV
            subtype ayw4, isolate hb321; isolate patient Ferracuti'83;
            isolate patient Castag'83; isolate patient Sanna'84;
            isolate patient Licheri'1'85; isolate patient Flore-1'86;
            isolate patient Licheri'83
            isolate patient 18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change
            10-Oct-1997
DATE
ACCESSIONS S47405; S53191; S53209; S53234; S53264; S53249; S53262;
            S53277; A03711
REFERENCE
AUTHORS     Plucienniczak, A.
SUBMISSION  submitted to the EMBL Data Library, August 1994
DESCRIPTION Molecular cloning and sequencing of two complete genomes of
            polish isolates of human hepatitis B virus.
            accession S47405
            ##molecule_type DNA
            ##residues 1-212 ##label PU
            ##cross-references EMBL:Z53716; NID:g527435; PID:g527437
            ##experimental_source subtype ayw4, isolate hb321
            S53112
REFERENCE
AUTHORS     Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
SUBMISSION  submitted to the EMBL Data Library, March 1995
DESCRIPTION isolate
            accession S53191
            ##molecule_type DNA
            ##residues 1-212 ##label LAI
            ##cross-references EMBL:X85283; NID:g736088; PID:g736090
            ##experimental_source isolate patient Ferracuti'83
            S53209
            ##molecule_type DNA
            ##residues 1-212 ##label LA2
            ##cross-references EMBL:X85280; NID:g736114; PID:g736116
            ##experimental_source isolate patient Castag'83
            S53234
            ##molecule_type DNA
            ##residues 1-212 ##label LA3
            ##cross-references EMBL:X85300; NID:g736150; PID:g736152
            ##experimental_source isolate patient Sanna'84
            S53264
            ##molecule_type DNA
            ##residues 1-212 ##label LA4
            ##cross-references EMBL:X85313; NID:g736194; PID:g736196

```

##experimental_source isolate patient Licheri-1'85
#accession S53249
##molecule_type DNA
##residues 1-212 ##label LA5
##cross-references EMBL:X85306; NID:q736172; PID:q736174
##experimental_source isolate patient Fiore-1'86
#accession S53262
##molecule_type DNA
##residues 1-212 ##label LA6
##cross-references EMBL:X85312; NID:q736191; PID:q736193
##experimental_source isolate patient Licheri'83
#accession S53277
##molecule_type DNA
##residues 30-212 ##label LA7
##cross-references EMBL:X85317; NID:q736211; PID:q736214
##experimental_source patient Giordo-2'86
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be processed

REFERENCE
#authors Ag3214
#journal Galibert, F.; Mandart, E.; Floussil, F.; Tiollais, P.;
Charnay, P.
#title Nature (1979) 281:646-650
Nucleotide sequence of the hepatitis B virus genome (subtype
ayw) in E. coli.
#cross-references MUID:81012091
#accession A03711
##molecule_type DNA
##residues 1-212 ##label GAL
##cross-references GB:J02203; NID:q329640; PID:q329642
##experimental_source subtype ayw

GENETICS
#gene C
#CLASSIFICATION #superfamily hepatitis B virus core antigen
#KEYWORDS alternative initiators; core protein
#FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product core antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY
#length 212 #molecular-weight 24350 #checksum 762

Query Match 62.4%; Score 1375; DB 1; Length 212;
Best Local Similarity 99.5%; Pred. No. 8.67e-201;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 88
QY 112 SMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLKFRQLMFWHISCLTFGRETVEIYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLKFRQLMFWHISCLTFGRETVEIYL 231
Db 149 VSEGVWIRTPPAYRPPNAPILSTPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSEGVWIRTPPAYRPPNAPILSTPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 2
ENTRY S53207 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
Castaa-2'87)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castaa-2'87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997 .

ACCESSIONS S53207
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53207
##molecule_type DNA
##residues 1-183 ##label LA1
##cross-references EMBL:X85289; NID:q736110; PID:q736113
##experimental_source isolate patient Castaa-2'87
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
#CLASSIFICATION #superfamily hepatitis B virus core antigen
#KEYWORDS core protein
#SUMMARY #length 183 #molecular-weight 21102 #checksum 2165

Query Match 62.3%; Score 1373; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 1.88e-200;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 60
QY 113 MDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 172
Db 61 CMGELMTLATWGVNLEDPASRDVSVYNTNGLKFRQLMFWHISCLTFGRETVEIYL 120
QY 173 CMGELMTLATWGVNLEDPASRDVSVYNTNGLKFRQLMFWHISCLTFGRETVEIYL 232
Db 121 SFGVWIRTPPAYRPPNAPILSTPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSRE 180
QY 233 SFGVWIRTPPAYRPPNAPILSTPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSRE 292
Db 181 SOC 183
QY 293 SOC 295

RESULT 3
ENTRY S53211 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Castag-1'85 and others)
ALTERNATE_NAMES HBC antigen; HBe antigen precursor / HBC antigen; pre-C/C
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate patient Castag-1'85; isolate patient Ferracuti-1'89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53211
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53211
##molecule_type DNA
##residues 1-212 ##label LA1
##cross-references EMBL:X85291; NID:q736117; PID:q736119
##experimental_source isolate patient Castag-1'85
#accession S53197
##molecule_type DNA
##residues 30-212 ##label LAW
##cross-references EMBL:X85284; NID:q736095; PID:q736098
##experimental_source isolate patient Ferracuti-1'89
#note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
#CLASSIFICATION #superfamily hepatitis B virus core antigen
#KEYWORDS alternative initiators; core protein
#FEATURE
1-29
30-212 #domain signal sequence #status predicted #label SIG\
30-178 #product core antigen #status predicted #label CAG\
#product e antigen #status predicted #label EAG\
30-178

179-212 #domain carboxyl-terminal propeptide #link EAG #status
 predicted #label ECP
SUMMARY #length 212 #molecular-weight 24336 #checksum 746

Query Match 62.3%; Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.28e-200;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 88
 |||||
Oy 112 SMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 171
 |||||

Db 89 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 148
 |||||
Oy 172 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 231
 |||||

Db 149 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
 |||||
Oy 232 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 291
 |||||

Db 209 ESQC 212
 |||||
Oy 292 ESQC 295

RESULT 4
ENTRY S32204 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
 (subtype ayw, patient C1000)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
 #subtype ayw, patient C1000
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
 08-Sep-1997

ACCESSIONS
REFERENCE S32204
KEYWORDS S32202
 Prestier-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.;
 Gerok, W.; Resenack, J.
 #submission submitted to the EMBL Data Library, March 1993
 #description identification and sequence analysis of hepatitis B virus DNA
 in immunological negative infection.
 #accession S32204
 #molecule-type DNA
 #residues 1-212 #label PRE
 #cross-references EMBL:X72702; NID:9288927; PID:9288930
 #experimental_source subtype ayw, patient C1000
 #note due to a stop codon between the alternative initiators
 the e antigen precursor cannot be produced

GENETICS
CLASSIFICATION C
KEYWORDS #superfamily hepatitis B virus core antigen
FEATURE 1-29 #domain signal sequence #status predicted #label SIG\
 30-212 #product core antigen #status predicted #label CAG\
 30-118 #product e antigen #status predicted #label EAG\
 179-212 #domain carboxyl-terminal propeptide #link EAG #status
 predicted #label ECP

SUMMARY #length 212 #molecular-weight 24336 #checksum 770

Query Match 62.3%; Score 1374; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.28e-200;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 88
 |||||
Oy 112 SMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 171
 |||||

Db 89 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 148
 |||||
Oy 172 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 231
 |||||

Db 149 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
 |||||
Oy 232 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 291
 |||||

Db 209 ESQC 212
 |||||
Oy 292 ESQC 295

Db 149 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
 |||||
Oy 232 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 291
 |||||

Db 209 ESQC 212
 |||||
Oy 292 ESQC 295

RESULT 5
ENTRY S20750 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
 (subtype ayw, patient C1)
ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus, HBV
 #subtype ayw, patient C1
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
 08-Sep-1997

ACCESSIONS
REFERENCE S20750
KEYWORDS S20745
 Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.;
 Portu, A.
 #submission submitted to the EMBL Data Library, March 1992
 #description Sequence analysis of HBV genomes isolated from patients with
 HBsAg negative chronic liver disease.
 #accession S20750
 #molecule-type DNA
 #residues 1-212 #label LAI
 #cross-references EMBL:X65258; NID:959434; PID:959436
 #experimental_source subtype ayw, patient C1

GENETICS
CLASSIFICATION C
KEYWORDS #superfamily hepatitis B virus core antigen
 alternative initiators; core protein
FEATURE 1-29 #domain signal sequence #status predicted #label SIG\
 30-212 #product core antigen #status predicted #label CAG\
 30-118 #product e antigen #status predicted #label EAG\
 179-212 #domain carboxyl-terminal propeptide #link EAG #status
 predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 147

Query Match 62.3%; Score 1373; DB 2; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.88e-200;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 88
 |||||
Oy 112 SMDIDPYKEFGATVLLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALRQAI 171
 |||||

Db 89 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 148
 |||||
Oy 172 LCMGELMTLATVGVNLEDPASRDLYSVYNTNMGKLFKROLMLFHSCLTGFRETIVIEL 231
 |||||

Db 149 VAFGWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 208
 |||||
Oy 232 VSGFWIRTPPYRPPNAPILSTLPETTVVRRGRSPRRRTPSPRRRSQSPRRRSQSR 291
 |||||

Db 209 ESQC 212
 |||||
Oy 292 ESQC 295

RESULT 6
ENTRY S53270 #type complete
TITLE core antigen - hepatitis B virus (isolate patient
 Licheri-2/87)
ALTERNATE_NAMES Hbc antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
 #isolate patient Licheri-2/87
 #variety isolate patient Licheri-2/87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

08-Sep-1997

ACCESSIONS S53270

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53270

##molecule-type DNA

##residues 1-183 ##label LAI

##cross-references EMBL:X85314; NID:g736201; PID:g736204

##experimental_source isolate patient Licheri-2787

##note due to a stop codon between the alternative initiators the e antigen precursor cannot be produced

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS core protein

SUMMARY #length 183 #molecular-weight 21102 #checksum 2199

Query Match 62.1%; Score 1370; DB 2; Length 183;

Best Local Similarity 98.4%; Pred. No. 5,98e-200;

Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 60

113 MDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 172

61 CMGDLMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 120

173 CMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 232

Db 121 SFGVWIRTPPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180

233 SFGVWIRTPPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 292

Db 181 SQC 183

293 SQC 295

QY

RESULT 7

ENTRY S53216 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen

CONTAINS core antigen; e antigen

ORGANISM #formal_name hepatitis B virus, HBV

#variety isolate patient Castag/3

DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53216

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53216

##molecule-type DNA

##residues 1-212 ##label LAI

##cross-references EMBL:X85293; NID:g736124; PID:g736126

##experimental_source isolate patient Castag/3

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS alternative initiators; core protein

FEATURE

1-29 #domain signal sequence #status predicted #label SIG\

30-212 #product core antigen #status predicted #label CAG\

30-178 #product e antigen #status predicted #label EAG\

179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24363 #checksum 752

Query Match 62.1%; Score 1370; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 5,98e-200;

Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 88

112 SMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 171

89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 148

172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 231

Db 149 VSRGWIPTPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208

232 VSRGWIPTPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 291

QY

Db 209 ESQC 212

292 ESQC 295

QY

RESULT 8

ENTRY S53272 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES HBe antigen precursor / Hbc antigen; pre-C/C antigen

CONTAINS core antigen; e antigen

ORGANISM #formal_name hepatitis B virus, HBV

#variety isolate patient Licheri-3790

DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53272

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53272

##molecule-type DNA

##residues 1-212 ##label LAI

##cross-references EMBL:X85315; NID:g736205; PID:g736207

##experimental_source isolate patient Licheri-3790

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS alternative initiators; core protein

FEATURE

1-29 #domain signal sequence #status predicted #label SIG\

30-212 #product core antigen #status predicted #label CAG\

30-178 #product e antigen #status predicted #label EAG\

179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24366 #checksum 446

Query Match 62.1%; Score 1370; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 5,98e-200;

Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 88

112 SMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTASALYREALSEPHSCPHHTALRQAI 171

89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 148

172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIELY 231

Db 149 VSRGWIPTPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208

232 VSRGWIPTPAVPPAPNPAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 291

QY

Db 209 ESQC 212

292 ESQC 295

QY

RESULT 9

ENTRY S53163 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES (isolate patient Vittorina/92)
CONTAINS Hbe antigen precursor / Hbc antigen; pre-C/C antigen
ORIGINISM core antigen; e antigen
#formal_name hepatitis B virus, HBV
#variety isolate patient Vittorina/92
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53163
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53163
##molecule_type DNA
##residues 1-212 #label LAI
##cross-references EMBL:X85256; NID:q736050; PID:q736052
##experimental_source isolate patient Vittorina/92

GENETICS C
#gene
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product e antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24322 #checksum 842

Query Match 62.1%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 8.80e-200;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 88
:|||||
Oy 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 171
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 148
:|||||
Oy 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 231
:|||||

Db 149 VSFGWIRTPPAYRPPAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
:|||||
Oy 232 VSFGWIRTPPAYRPPAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 291
:|||||

Db 209 ESOC 212
:|||||
Oy 292 ESOC 295

RESULT 10
ENTRY S20746 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C)

ALTERNATE_NAMES Hbe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORIGINISM #formal_name hepatitis B virus, HBV
#variety subtype ayw, patient C
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change
08-Sep-1997

ACCESSIONS S20746
REFERENCE S20745
#authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.;
Porru, A.
#submission submitted to the EMBL Data Library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with
HBsAg negative chronic liver disease.
#accession S20746
##molecule_type DNA
##residues 1-212 #label LAI
##cross-references EMBL:X65257; NID:q59429; PID:q59431
##experimental_source subtype ayw, patient C

GENETICS C
#gene

CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product e antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 506

Query Match 62.1%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 8.80e-200;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 88
:|||||
Oy 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 171
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 148
:|||||
Oy 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 231
:|||||

Db 149 VSFGWIRTPPAYRPPAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 208
:|||||
Oy 232 VSFGWIRTPPAYRPPAPILSTLPTTVVRRGRSPRRRTPSPRRRSQSR 291
:|||||

Db 209 ESOC 212
:|||||
Oy 292 ESOC 295

RESULT 11
ENTRY S53225 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(isolate patient Chigline-2/86)

ALTERNATE_NAMES Hbe antigen precursor / Hbc antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORIGINISM #formal_name hepatitis B virus, HBV
#variety isolate patient Chigline-2/86
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change
08-Sep-1997

ACCESSIONS S53225
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53225
##molecule_type DNA
##residues 1-212 #label LAI
##cross-references EMBL:X85296; NID:q736137; PID:q736139
##experimental_source isolate patient Chigline-2/86

GENETICS C
#gene
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-212 #product e antigen #status predicted #label CAG\
179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP

SUMMARY #length 212 #molecular-weight 24308 #checksum 655

Query Match 62.1%; Score 1369; DB 2; Length 212;
Best Local Similarity 98.4%; Pred. No. 8.80e-200;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 88
:|||||
Oy 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALESPHCSPHNTALQAI 171
:|||||

Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 148
:|||||
Oy 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNGLKFRQLMFHISCLFGRETVEYL 231
:|||||

Db	149	VSFGWIRTPPYRPPNAPILSTLEETIVRRGRGSPRRRPSPPRRRSQSPRRRSQSR	208
Qy	232	VSFGWIRTPPYRPPNAPILSTLEETIVRRGRGSPRRRPSPPRRRSQSPRRRSQSR	291
Db	209	ESQC 212	
Qy	292	ESQC 295	
RESULT	12		
ENTRY			
TITLE			
ALTERNATE_NAMES			
CONTAINS			
ORGANISM			
DATE			
ACCESSIONS			
REFERENCE			
#authors			
#submission			
#accession			
#molecule_type			
#residues			
#cros-references			
#experimental_source			
GENETICS			
#gene			
CLASSIFICATION			
KEYWORDS			
FEATURE			
1-29			
30-212			
30-178			
179-212			
SUMMARY			
Query Match			
Best Local Similarity			
Matches			
Db	29	GMDIDPYKEFGATVLLSFLPSAFPPSVRDLLDTASATYREALSPHCSPHHTALRQAI	88
Qy	112	SMDDIPYKEFGATVLLSLPSDFPSVRDLDTASATYREALSPHCSPHHTALRQAI	171
Db	89	LCMGDLMTLATWGVNLEDPASRDLYVSVYNTNMGLKTRQLLMTFHSICLTFGRETIVIEL	148
Qy	172	LCMGDLMTLATWGVNLEDPASRDLYVSVYNTNMGLKTRQLLMTFHSICLTFGRETIVIEL	231
Db	149	VSFGWIRTPPYRPPNAPILSTLEETIVRRGRGSPRRRPSPPRRRSQSPRRRSQSR	208
Qy	232	VSFGWIRTPPYRPPNAPILSTLEETIVRRGRGSPRRRPSPPRRRSQSPRRRSQSR	291
Db	209	ESQC 212	
Qy	292	ESQC 295	
RESULT	13		
ENTRY			
TITLE			
ALTERNATE_NAMES			
CONTAINS			
ORGANISM			
DATE			
ACCESSIONS			

REFERENCE	553112	Lai, M.E.; Mazzeoli, A.P.; Porri, A.; Balestrieri, A.
#authors		
#submission		submitted to the EMBL Data Library, March 1995
#accession	S53223	
##molecule_type	DNA	
##residues	1-212	#label LAI
##cross-References	EMBL:X85325; NID:g736134; PID:g736136	
##experimental_source	isolate	patient Chighine-1-85
GENETICS		
#gene	C	#superfamily hepatitis B virus core antigen
CLASSIFICATION		alternative initiators; core protein
KEYWORDS		
FEATURE	1-29	
30-212		#domain signal sequence #status predicted #label SIG\
30-178		#product core antigen #status predicted #label CAG\
179-212		#product e antigen #status predicted #label EAG\ #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP
SUMMARY	#length 212	#molecular-weight 24364 #checksum 1123
Query Match	62.0%; Score 1366; DB 2; Length 212;	
Best Local Similarity	97.8%; Pred. No. 2,806-159;	
Matches	180; Conservative 3; Mismatches 1; Indels 0; Gaps 0	
Db	29 GMDIDPKYEFQATVLLSFLPSDFPSRDLDTASALYREALESPHCSPHNTALROAI	88
Qy	112 SMDIDPKYEFQATVLLSFLPSDFPSRDLDTASALYREALESPHCSPHNTALROAI	171
Db	89 LCMGDLMSLATVAGNLEDPISRDLYSVYVNTNMGKFRQLAMFISCTFGRETVIEL	148
Qy	172 LCMGDLMSLATVAGNLEDPISRDLYSVYVNTNMGKFRQLAMFISCTFGRETVIEL	231
Db	149 VSEGWIRTPPAYRPPNAPILSTLPEITTVRRRGSPPRRTPSPRRRSQSPPRRRSQSR	208
Qy	233 VSEGWIRTPPAYRPPNAPILSTLPEITTVRRRGSPPRRTPSPRRRSQSPPRRRSQSR	291
Db	209 E50C 212	
Qy	292 E50C 295	
RESULT	14	
ENTRY	S53274	#type complete
TITLE	e antigen precursor / core antigen - hepatitis B virus	
ALTERNATE_NAMES	(isolate patient Giordo/84)	
CONTAINS	HBE antigen precursor / HBC antigen; pre-C/C antigen	
ORGANISM	core antigen; e antigen	
variety	#formal_name hepatitis B virus, HBV	
DATE	isolate patient Giordo/84	
08-Jul-1995	#sequence_revision 03-Aug-1995	#text_change
08-Sep-1997		
ACCESSIONS	S53274	
REFERENCE	S53112	
#authors	Lai, M.E.; Mazzeoli, A.P.; Porri, A.; Balestrieri, A.	
#submission	submitted to the EMBL Data Library, March 1995	
#accession	S53274	
##molecule_type	DNA	
##residues	1-212	#label LAI
##cross-References	EMBL:X85316; NID:g736208; PID:g736210	
##experimental_source	isolate	patient Giordo/84
GENETICS		
#gene	C	#superfamily hepatitis B virus core antigen
CLASSIFICATION		alternative initiators; core protein
KEYWORDS		
FEATURE	1-29	
30-212		#domain signal sequence #status predicted #label SIG\
30-178		#product core antigen #status predicted #label CAG\ #product e antigen #status predicted #label EAG\ #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP
179-212		
SUMMARY	#length 212	#molecular-weight 24376 #checksum 390
Query Match	62.0%; Score 1366; DB 2; Length 212;	

Best Local Similarity 98.9%; Pred. No. 2,80e-199;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSEPHHCSPHHTALROAI 88
:|||||
QY 112 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSEPHHCSPHHTALROAI 171
:|||||
Db 89 LCMGELMTLATWGVNLEDPSARDLVSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 148
:|||||
QY 172 LCMGELMTLATWGVNLEDPSARDLVSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 231
:|||||
Db 149 VSGVWIRTPPAYRPPNAPILTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 208
:|||||
QY 232 VSGVWIRTPPAYRPPNAPILTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 291
:|||||
Db 209 ESOC 212
:|||||
QY 292 ESOC 295

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu/89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#variety isolate patient Muresu/89
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53169
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
##molecule_type DNA
##residues 1-183 ##label LAI
##cross_references EMBL:X85275; NID:g736057; PID:g736060
##experimental_source isolate patient Muresu/89
##note due to a stop codon between the alternative initiators
the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular_weight 21088 #checksum 1849

Query Match 61.9%; Score 1364; DB 2; Length 183;
Best Local Similarity 98.4%; Pred. No. 6.07e-199;
Matches 180; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSEPHHCSPHHTALROAIL 60
:|||||
QY 113 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSEPHHCSPHHTALROAIL 172
:|||||
Db 61 CWMGELMTLATWGVNLEDPSARDLVSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 120
:|||||
QY 173 CWMGELMTLATWGVNLEDPSARDLVSVYNTNMGKFRQLMFHISCLTFGRETIVIEYL 232
:|||||
Db 121 SFGVWIRTPPAYRPPNAPILTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 180
:|||||
QY 233 SFGVWIRTPPAYRPPNAPILTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 292
:|||||
Db 181 SOC 183
:|||||
QY 293 SOC 295

Search completed: Thu Dec 16 13:49:08 1999
Job time : 74 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSPch_DP protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:49:26 1999; MasPar time 10.05 Seconds
Tabular output not generated. 829,670 Million cell updates/sec

Title: >US3890-10-38183
Description: (1-295) from US3890-10-38183.pep
Perfect Score: 2205
Sequence: 1 MDIDPYKEGATVELLSFLP.....RRRRSSPPRRRSQSRESOC 295

Scoring table: PAM 150
Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot37
1:swissprot

Statistics: Mean 48.376; Variance 105.353; scale 0.459

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1374	62.3	183	1	CORA_HPBV CORE ANTIGEN.	2,04e-230
2	1358	61.6	211	1	CORA_HPBVA CORE ANTIGEN.	2,47e-227
3	1349	61.2	183	1	CORA_HPBV2 CORE ANTIGEN.	1,34e-225
4	1336	60.6	183	1	CORA_HPBVO CORE ANTIGEN.	4,26e-223
5	1334	60.5	183	1	CORA_HPBV4 CORE ANTIGEN.	1,03e-222
6	1332	60.0	183	1	CORA_HPBV1 CORE ANTIGEN.	2,11e-220
7	1321	59.9	185	1	CORA_HPBV CORE ANTIGEN.	3,29e-220
8	1314	59.6	195	1	CORA_HPBV CORE ANTIGEN.	7,31e-219
9	1313	59.5	212	1	CORA_HPBV CORE ANTIGEN.	1,14e-218
10	1308	59.3	183	1	CORA_HPBVL CORE ANTIGEN.	1,04e-217
11	1305	59.2	185	1	CORA_HPBV2 CORE ANTIGEN.	3,94e-217
12	1304	59.1	214	1	CORA_HPBV CORE ANTIGEN.	6,14e-217
13	938	42.5	188	1	CORA_HBV1 CORE ANTIGEN.	5,57e-147
14	934	42.4	217	1	CORA_HBVS CORE ANTIGEN.	3,20e-146
15	933	42.3	187	1	CORA_HBV8 CORE ANTIGEN.	4,94e-146
16	200	9.1	305	1	CORA_HPBHE CORE ANTIGEN.	5,88e-14
17	200	9.1	305	1	CORA_HPBHE CORE ANTIGEN.	5,88e-14
18	200	9.1	305	1	CORA_HPBDC CORE ANTIGEN.	5,88e-14
19	200	9.1	305	1	CORA_HPBDC CORE ANTIGEN.	5,88e-14
20	200	9.1	305	1	CORA_HPBDC CORE ANTIGEN.	5,88e-14
21	187	8.5	289	1	CERP_HESAU CORE ANTIGEN.	5,64e-12
22	185	8.4	493	1	CERP_HUMAN CHOLESTERYL ESTER TRAN	1,13e-11
23	179	8.1	493	1	CERP_MACFA CHOLESTERYL ESTER TRAN	8,92e-11

24	168	7.6	497	1	CERP_RABIT CHOLESTERYL ESTER TRAN	3,71e-09
25	158	7.2	196	1	SPR2_CAEL PUTATIVE SPLICING FACT	1,02e-07
26	148	6.7	208	1	YX22_CAEL HYPOTHETICAL 24.0 KD P	2,58e-06
27	136	6.2	78	1	PR1_SEPOF SPERMATID-SPECIFIC PRO	1,10e-04
28	131	5.9	61	1	HSP1_MACD SPERMATID-SPECIFIC PRO	5,04e-04
29	130	5.9	77	1	PR2_SEPOF SPERMATID-SPECIFIC PRO	6,80e-04
30	125	5.7	60	1	HSP1_MACI SPERMATID-SPECIFIC PRO	3,00e-03
31	125	5.7	61	1	HSP1_MACRG SPERMATID-SPECIFIC PRO	3,00e-03
32	126	5.7	498	1	VE2_HV08 REGULATORY PROTEIN E2	2,24e-03
33	126	5.7	1523	1	SON_HUMAN SON PROTEIN (SON3)	2,24e-03
34	124	5.6	57	1	HSP1_DIDMA SPERMATID-SPECIFIC PRO	4,02e-03
35	121	5.5	61	1	HSP1_CHICK SPERMATID-SPECIFIC PRO	9,62e-03
36	122	5.5	91	1	PR1_MYED SPERMATID-SPECIFIC PRO	7,20e-03
37	122	5.5	132	1	PR1_ANTGR SPERMATID-SPECIFIC PRO	7,20e-03
38	121	5.5	739	1	DD13_CAEL PUTATIVE PRE-MRNA SPLI	9,62e-03
39	121	5.5	843	1	CY1_BRUA PEPTIDYLISOMERASE	9,62e-03
40	119	5.4	50	1	HSP1_RAT SPERMATID-SPECIFIC PRO	1,71e-02
41	119	5.4	50	1	HSP1_MOUSE SPERMATID-SPECIFIC PRO	1,71e-02
42	118	5.4	238	1	SER7_HUMAN SPERMATID-SPECIFIC PRO	2,27e-02
43	118	5.4	576	1	CATA_RHOCA PEROXIDASE / CATALASE	2,27e-02
44	116	5.3	492	1	PEPD_MOUSE XAA-PRO DIPEPTIDASE (E	4,00e-02
45	115	5.2	103	1	HSP2_MACNE SPERMATID-SPECIFIC PRO	5,30e-02

ALIGNMENTS

RESULT 1
ID CORA_HPBV STANDARD; PRT; 183 AA.

AC P03146:
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.

OS HEPATITIS B VIRUS (SUBTYPE AYV).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN SEQUENCE FROM N.A.

RA GALBERT F., MANDART E., FITOUSSI F., TIOLLAIS P., CHARNEY P.

RT "Nucleotide sequence of the hepatitis B virus genome (subtype ayw)

RL cloned in E. coli." (1979).

[2]

RP SEQUENCE FROM N.A. (CLONE PHB320).

RX MEDLINE: 85204397.

RA BICHKO V., PUSHKO P., DRELLINA D., PUMPER P., GREN E.

RT "Subtype ayw variant of hepatitis B virus. DNA primary structure

RT analysis." (1985).

RL FEBS LETT. 185:208-212(1985).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
The European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

DR EMBL: V01460; G62278; ALT_INT.

DR EMBL: X02496; NOT_ANNOTATED_CDS.

DR PIR: A03711; NIKVLBH.

DR PIR: A03712; NIKVLBH.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN; REPEAT.

FT REPEAT 162 169

FT REPEAT 170 177

FT CONFLICT 33 33 T -> N (IN REF. 2).

FT CONFLICT 80 80 A -> I (IN REF. 2).

SO SEQUENCE 183 AA; 2116 MW; 2AE7A17A CR32;

Query Match 62.3%; Score 1374; DB 1; Length 183;
Best local similarity 100.0%; Pred. No. 2,04e-230;

Matches 183; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKEFGATVELLSFLPSDFPVSRLDITASALYREALSPHCSPHHTAL 60
AC P24023;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DE 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN HEPATITIS B VIRUS (STRAIN ALPHA1).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90266476.
RA TONG S., LI J., VIVITSKI L., TREPO C.;
RT Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M32138; G495034; ALT_SEQ.
DR PIR; A34773; NKVLAI.
DR PFAM; PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 183 190
FT REPEAT 198 206
SQ SEQUENCE 211 AA; 24208 MW; 0924CE5E CRC32;

Query Match 61.6%; Score 1358; DB 1; Length 211;
Best Local Similarity 95.2%; Pred. No. 2.47e-227;
Matches 179; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 24 LGWMDMDIDPKEFGATVELLSFLPSDFPVSRLDITASALYREALSPHCSPHHTAL 83
AC P17392;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DE 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN OKINAWA/PODW287).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 8910694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
RA MIYAKAWA Y., MAYUMI M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes.";
RL
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; J02202; G329638; -
DR EMBL; A08967; G411874; -
DR PIR; B93217; NKVLAI2.
DR PFAM; PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;

Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPKEFGATVELLSFLPSDFPVSRLDITASALYREALSPHCSPHHTAL 60
AC P24023;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DE 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN HEPATITIS B VIRUS (STRAIN ALPHA1).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90266476.
RA TONG S., LI J., VIVITSKI L., TREPO C.;
RT Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; J02202; G329638; -
DR EMBL; A08967; G411874; -
DR PIR; B93217; NKVLAI2.
DR PFAM; PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;

Query Match 61.2%; Score 1349; DB 1; Length 183;
Best Local Similarity 96.7%; Pred. No. 1.34e-225;
Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPKEFGATVELLSFLPSDFPVSRLDITASALYREALSPHCSPHHTAL 60
AC P24023;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DE 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN HEPATITIS B VIRUS (STRAIN ALPHA1).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90266476.
RA TONG S., LI J., VIVITSKI L., TREPO C.;
RT Active hepatitis B virus replication in the presence of anti-HBe is associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; J02202; G329638; -
DR EMBL; A08967; G411874; -
DR PIR; B93217; NKVLAI2.
DR PFAM; PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21042 MW; 6EA6B674 CRC32;


```

RL J. GEN. VIROL. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its use
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00330; ?; NOT ANNOTATED_CDS.
DR PIR; B28925; NKVLJ2.
DR PFAM; PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA; 21081 MW; 115BD9E3 CRC32;

Query Match 60.6%; Score 1336; DB 1; Length 183;
Best Local Similarity 96.2%; Pred. No. 4,26e-23;
Matches 176; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1 MDIDYKEGASVELLSFLPSDFPSVDRDLTASALYREALSEPHCSPHHTALROATL 60
Oy 113 MDIDYKEGATVELLSFLPSDFPSVDRDLTASALYREALSEPHCSPHHTALROATL 172
Db 61 CWSGLMNLATWGSNLEDPASRELVSYYNVMGLKIRQLMFHISCLTFGRETVEYLV 120
Oy 173 CWSGLMNLATWGSNLEDPASRDVLVSYNTNMGLKFRQLMLFHSCLTFGRETVEYLV 232
Db 121 SFGVWIRPPAPRPAPNAPLSTLPETTVYRRGRSPRRTPSPRRRRSSPPRRRSQSE 180
Oy 233 SFGVWIRPPAPRPAPNAPLSTLPETTVYRRGRSPRRTPSPRRRRSSPPRRRSQSE 292
Db 181 SOC 183
Oy 293 SOC 295

RESULT 5 STANDARD: PRT: 183 AA.
AC P03150; P03151;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADR), HEPATITIS B VIRUS (SUBTYPE ADR4),
OS AND HEPATITIS B VIRUS (SUBTYPE ADM / STRAIN INDOONESIA/PIDM420).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ADR:
RX MEDLINE; 83168919.
RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
RT DNA; subtype adr and adm.";
RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ADR4:
RX MEDLINE; 83246570.
RA FUJIIYAMA A., MIYANOHARA A., NOZAKI C., YONEYAMA T., OHTOMO N.,
RA MATSUURA K.;
RT "Cloning and structural analyses of hepatitis B virus DNAs, subtype
RT adm.";
RL NUCLEIC ACIDS RES. 11:4601-4610(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ADM:
RX MEDLINE; 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAMA H., SASTROSOMIGNIO R.I., IMAI M.,
RA MIYAKAWA Y., MATSUMI M.;

```

```

RT      "Typing hepatitis B virus by homology in nucleotide sequence:
RL      Comparison of surface antigen subtypes.";
RL      J. GEN. VIROL. 69:2575-2583(1988).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; V00867; -; NOT_ANNOTATED_CDS.
DR      EMBL; X01867; G59407; -;
DR      EMBL; D00331; -; NOT_ANNOTATED_CDS.
DR      PIR; A93480; NKVL45.
DR      PIR; B93460; NKVL44.
DR      PIR; C28925; NKVLJ3.
DR      PRAM; PF00906; Hepatitis_core; 1.
KW      CORE PROTEIN; REPEAT.
FT      REPEAT      162      169
FT      REPEAT      170      177
SQ      SEQUENCE   183 AA;  21095 MW;  D2E185B9 CRC32;
Cc      -----
Db      Query Match      60.5%; Score 1334; DB 1; Length 183;
Cc      Best Local Similarity 95.6%; Pred. No. 1,036-222;
Cc      Matches 175; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Cc      -----
Db      1 MDIDYKEFGASVELSLSPDSDFPSIRDLDTASALYREALSEPHCSPHHTARQAIL 60
Cc      |||||
Cc      113 MDIDYKEFGATVELSLSPDSDFPSVRDLDTASALYREALSEPHCSPHHTARQAIL 172
Cc      -----
Db      61 CWGELMLNATWGSNLSELPASRELIVSYVNNMGKIRQLLMFHISCTFGRETVLEYLY 120
Cc      |||||
Cc      173 CWGELMLNATWGVNLSELPASRDLVSYVNNMGKIFRQLLMFHISCTFGRETVLEYLY 232
Cc      -----
Db      121 SFGWIRTPPAYRPNPADILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQGRE 180
Cc      |||||
Cc      233 SFGWIRTPPAYRPNPADILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQGRE 292
Cc      -----
Db      181 SQC 183
Cc      |||
Cc      293 SQC 295
Cc      -----
Cc      RESULT      6      STANDARD;      PRI:      183 AA.
Cc      ID      CORA_HPBVJ      STANDARD;      PRI:      183 AA.
Cc      AC      P17391;
Cc      DT      01-AUG-1990 (REL. 15, CREATED)
Cc      DT      01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
Cc      DT      01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
Cc      DE      CORE ANTIGEN.
Cc      OS      HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN JAPAN/BJDM233).
Cc      OS      VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
Cc      [1]
Cc      SEQUENCE FROM N.A.
Cc      RX      MEDLINE: 89010694.
Cc      RA      OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
Cc      RA      MIYAKAWA Y., MATSUMI M.,
Cc      RT      "Typing hepatitis B virus by homology in nucleotide sequence:
Cc      RT      comparison of surface antigen subtypes.";
Cc      J. GEN. VIROL. 69:2575-2583(1988).
Cc      -----
Cc      This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc      between the Swiss Institute of Bioinformatics and the EMBL outstation -
Cc      the European Bioinformatics Institute. There are no restrictions on its
Cc      use by non-profit institutions as long as its content is in no way
Cc      modified and this statement is not removed. Usage by and for commercial
Cc      entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc      or send an email to license@isb-sib.ch).
Cc      -----
Cc      EMBL; D00329; -; NOT_ANNOTATED_CDS.
Cc      PIR; A28925; NKVLJ1.

```

DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN: REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 21224 MW: F6B348B6 CRC32:
Query Match 60.0%; Score 1322; DB 1; Length 183;
Best Local Similarity 96.2%; Pred. No. 2,11e-220;
Matches 176; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROAIL 60
QY 113 MDIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROAIL 172
Db 61 CWGELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEYLY 120
QY 173 CWGELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEYLY 232
Db 121 SFGVWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTPSPRRRSOSPRE 180
QY 233 SFGVWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTPSPRRRSOSPRE 292
Db 181 SOC 183
QY 293 SOC 295

RESULT 7
ID CORA_HPBVT STANDARD; PRT; 185 AA.
AC P03149;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 83168919.
RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA; subtype adr and adw.";
RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL: V00866; -- NOT_ANNOTATED_CDS.
DR PIR: C93460; NKVLA6.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN: REPEAT.
FT REPEAT 162 171
FT REPEAT 172 179
SQ SEQUENCE 185 AA: 21394 MW: 791E0381 CRC32:
Query Match 59.9%; Score 1321; DB 1; Length 185;
Best Local Similarity 96.2%; Pred. No. 3,29e-220;
Matches 178; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Db 1 MDIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROAIL 60
QY 113 MDIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROAIL 172
Db 61 CWGELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEYLY 120
QY 173 CWGELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEYLY 232

Db 121 SFGVWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTPSPRRRSOSPRE 180
QY 233 SFGVWIRTPPAYRPNPAPILSTLPETTVVRR--GSPRRRTPSPRRRSOSPRE 290
Db 181 RESOC 185
QY 291 RESOC 295

RESULT 8
ID CORA_HPBVT STANDARD; PRT; 195 AA.
AC P29178;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 90169850.
RA BHAT R.A., ULRICH P.P., VYAS G.N.;
RT "Molecular characterization of a new variant of hepatitis B virus in
a persistently infected homosexual man.";
RL HEPATOLOGY 11:271-276(1990).
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN: REPEAT.
FT REPEAT 174 181
FT REPEAT 182 189
SQ SEQUENCE 195 AA: 22461 MW: AF3DB5F3 CRC32:
Query Match 59.6%; Score 1314; DB 1; Length 195;
Best Local Similarity 93.5%; Pred. No. 7,31e-219;
Matches 173; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 11 FGDLIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROA 70
QY 111 LMSDIDPYKEGATVELLSFSDPFPSVRLDLPASALYREALSPHCSPHHTALROA 170
Db 71 ILGWELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEY 130
QY 171 ILGWELMTLATWGNLDPASRDLYVYVNTNMGLKITROLMEFHISCLTFGRVLEY 230
Db 131 LVSEGVWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTPSPRRRSOSPRE 190
QY 231 LVSEGVWIRTPPAYRPNPAPILSTLPETTVVRRGRSPRRRTPSPRRRSOSPRE 290
Db 191 RESOC 195
QY 291 RESOC 295

RESULT 9
ID CORA_HPBVT STANDARD; PRT; 212 AA.
AC O05495;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE W4B).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 93346870.
RA NAUMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,
RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
that expresses HBV surface antigen subtype adw4.";
RL J. GEN. VIROL. 74:1627-1632(1993).
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69798; G59423; -.
DR PFM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT DOMAIN 178 204 ARG-RICH.
FT DOMAIN 184 203 3 X 5 AA REPEATS OF S-P-R-R-R.
FT REPEAT 184 188 1.
FT REPEAT 191 195 2.
FT REPEAT 199 203 3.
SO SEQUENCE 212 AA; 24234 MW; BCCDF263 CRC32;
Query Match 59.5%; Score 1313; DB 1; Length 212;
Best Local Similarity 92.9%; Pred. No. 1,14e-218;
Matches 171; Conservative 10; Mismatches 3; Indels 0; Gaps 0;
Dd 29 GMDIDYKEFGASVELLSFLPSDFPSPVRDLDTLSALYRALESPHECTPNHTALRQAI 88
Dy 112 SMDIDPYKEFGATVELLSFLPSDFPSPVRDLDTLSALYREALSPHECTPNHTALRQAI 171
Dd 89 LCMGELMLTASVGNLDDPARDLVYVYNTNMGCLKRQLMFIISCLTGREVLEYL 148
Dy 172 LCMGELMLTATVGNLDDPASRDVSVYVNTNMGCLKRQLMFIISCLTGREVLEYL 231
Dd 149 VSFQWIRTPPAYRPPNPILSTLPETVVVRGRSPRRRPPSPRRRSOSP RRRRSOSP 208
Dy 232 VSFQWIRTPPAYRPPNPILSTLPETVVVRGRSPRRRPPSPRRRSOSP RRRRSOSP 291
Dd 209 ASQC 212
Dy 292 ESQC 295
RESULT 10
ID CORA_HPBVL STANDARD: PRT: 183 AA.
AC P12901;
DT 01-OCT-1989 (REL. 12, CREATED)
DT 01-OCT-1989 (REL. 13, LAST SEQUENCE UPDATE)
DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN LSH / CHIMPANZEE ISOLATE).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88258473.
RA VAUDIN M., WOLSTENHOLME A.J., TSIOUAYE K.N., ZUCKERMAN A.J.,
RA HARRISON T.J.;
RT "The complete nucleotide sequence of the genome of a hepatitis B
RT virus isolated from a naturally infected chimpanzee.";
RL J. GEN. VIROL. 65:1383-1385(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00220; D1000603; -.
DR PIR: A28885; NKVLCF.
DR PFM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SO SEQUENCE 183 AA; 20999 MW; 1AEF57C9 CRC32;

```

Query Match	59.3%	Score 1308:	DB 1:	Length 183:
Best Local Similarity	94.5%	Pred. No. 1.04e-217:		
Matches 173:	Conservative	6:	Mismatches 4:	Indels 0: Gaps 0:
Db	1	MDIDPYKEFGATVLLSFLPSDFEPPSVRDLLDTASALYREALSPHCSPHNHALRQAIL	60	
Qy	113	MDIDPYKEFGATVLLSFLPSDFEPPSVRDLLDTASALYREALSPHCSPHNHALRQAIL	172	
Db	61	CWGLMTLATVGNLDDPASRDLYVSVYNTNGKLKROLMLFHSICLTGTGRETVLEYLV	120	
Qy	173	CWGLMTLATVGNLDDPASRDLYVSVYNTNGKLKROLMLFHSICLTGTGRETVLEYLV	232	
Db	121	SEGVWIRTPPAPRPNPAILSTLPETTVVRRGRSPRRTPSPRRRSQSPRRRRSQSPA	180	
Qy	233	SEGVWIRTPPAPRPNPAILSTLPETTVVRRGRSPRRTPSPRRRSQSPRRRRSQSRE	292	
Db	181	SQC 183		
Qy	293	SQC 295		
RESULT	11			
ID	CORA_HPBV2	STANDARD;	PRT;	185 AA.
AC	P03148;			
DT	21-JUL-1986 (REL. 01, CREATED)			
DT	01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)			
DT	15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			
OC	HEPATITIS B VIRUS (SUBTYPE ADW2).			
OS	VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.			
RP	SEQUENCE FROM N.A.			
RN	[1]			
RA	VALENZUELA P., QUIROGA M., ZALDIVAR J., GRAY P., RUTTER W.J.;			
RL	(IN) FIELD B.N., JAKENISCH R., FOX C.F. (EDS.);			
RL	ANIMAL VIRUS GENETICS, PP.57-70, ACADEMIC PRESS, NEW YORK (1980).			
DR	PIR; A94409; NKLVA3.			
DR	PFAM; PF00906; Hepatitis_core; 1.			
KW	CORE PROTEIN; REPEAT.			
FT	REPEAT 164 171			
FT	REPEAT 172 179			
SO	SEQUENCE 185 AA; E2EA3360 CIRC32:			
Query Match	59.2%	Score 1305:	DB 1:	Length 185:
Best Local Similarity	95.7%	Pred. No. 3.94e-217:		
Matches 177:	Conservative	4:	Mismatches 2:	Indels 2: Gaps 1:
Db	1	MDIDPYKEFGATVLLSFLPSDFEPPSVRDLLDTASALYREALSPHCSPHNHALRQAIL	60	
Qy	113	MDIDPYKEFGATVLLSFLPSDFEPPSVRDLLDTASALYREALSPHCSPHNHALRQAIL	172	
Db	61	CWGLMTLATVGNLDDPASRDLYVSVYNTNGKLKROLMLFHSICLTGTGRETVLEYLV	120	
Qy	173	CWGLMTLATVGNLDDPASRDLYVSVYNTNGKLKROLMLFHSICLTGTGRETVLEYLV	232	
Db	121	SEGVWIRTPPAPRPNPAILSTLPETTVVRRGRSPRRTPSPRRRSQSPRRRRSQS	180	
Qy	233	SEGVWIRTPPAPRPNPAILSTLPETTVVRR--GRSPRRTPSPRRRSQSPRRRRSQS	290	
Db	181	RESQC 185		
Qy	291	RESQC 295		
RESULT	12			
ID	CORA_HPBV9	STANDARD;	PRT;	214 AA.
AC	P17099;			
DT	01-AUG-1990 (REL. 15, CREATED)			
DT	01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)			
DT	01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)			
DE	CORE ANTIGEN.			
GN	C.			

```

OS HEPATITIS B VIRUS (SUBTYPE ADV / STRAIN 991).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX KOEHEL H.G., SCHUELER A., LOTTMANN S., THOMSEN R.;
RA SUBMITTED (FEB-1990) TO EMBL/GENEBANK/DBJ DATA BANKS.
RL
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X51970; G60433; -.
DR PIR: S10381; NKVLKS.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 193 200
FT REPEAT 201 208
SQ SEQUENCE 214 AA: 24722 MW: 786CD048 CRC32;

Query Match 59.1%; Score 1304; DB 1; Length 214;
Best Local Similarity 95.2%; Pred. No. 6.14e-217;
Matches 177; Conservative 5; Mismatches 2; Indels 2; Gaps 1;

Db 29 GMDIDPKYKFGATVLLSPDPSFVSVDLDTASALREALSEPHHTRALROI 88
QY :|||||
112 SMIDPKYKFGATVLLSPDPSFVSVDLDTASALREALSEPHHTRALROI 171
Db 89 LCMGELMTLATWGNLLEDPASRDVAVYNTNMGLKIRQLMFRISYLTFGREIVLEYL 148
QY :|||||
172 LCMGELMTLATWGNLLEDPASRDVAVYNTNMGLKIRQLMFRISYLTFGREIVLEYL 231
Db 149 VSPGWIRPPAPYRPPNAPILSTLPETTYVRRDRGRSPRRRSPRRRSQ 208
QY :|||||
232 VSPGWIRPPAPYRPPNAPILSTLPETTYVRRR--GRSPRRRTPSPRRRSQ 289
Db 209 SRESQC 214
QY :|||||
290 SRESQC 295

RESULT 13
ID CORA_MHV1 STANDARD: PRT; 188 AA.
AC P03152;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 1 (MHV 1),
OS WOODCHUCK HEPATITIS VIRUS 7 (MHV 7),
OS WOODCHUCK HEPATITIS VIRUS 59 (MHV 59), AND
OS WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE) (MHV 8).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=WOODCHUCK HEPATITIS VIRUS 1;
RX MEDLINE: 82216969.
RA GALIBERT F., CHEN T.N., MANDART E.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT comparison with the hepatitis B virus sequence.";
RL J. VIROL. 41:51-65(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=WOODCHUCK HEPATITIS VIRUS 7, AND WOODCHUCK HEPATITIS VIRUS 59;
RX MEDLINE: 88101359.
RA COHEN J.I., MILLER R.H., ROSENBLUM B., DENNISTON K., GERIN J.L.,
RA PURELL R.H.;
RT "Sequence comparison of woodchuck hepatitis virus replicative forms
RT shows conservation of the genome.";

```

```

RL VIROLOGY 162:12-20(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=WOODCHUCK HEPATITIS VIRUS 8 (INFECTIOUS CLONE);
RX MEDLINE: 89184524.
RA GIRONES R., COTE P.J., HORNBuckle W.E., TENNANT B.C., GERIN J.L.,
RA PURELL R.H., MILLER R.H.;
RT "Complete nucleotide sequence of a molecular clone of woodchuck
RT hepatitis virus that is infectious in the natural host.";
RL PROC. NATL. ACADE. SCI. U.S.A. 86:1846-1849(1989).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J02442; G336129; -.
DR EMBL: M18752; G336140; -.
DR EMBL: M19183; G336145; -.
DR EMBL: J04514; G336149; -.
DR PIR: A03713; NKVLC.
DR PIR: C32397; NKVLC.
DR PIR: C32397; NKVLC4.
DR PIR: C32397; NKVLC3.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 167 174
FT REPEAT 175 182
SQ SEQUENCE 188 AA: 21693 MW: A667DB27 CRC32;

Query Match 42.5%; Score 938; DB 1; Length 188;
Best Local Similarity 67.0%; Pred. No. 5.57e-147;
Matches 126; Conservative 25; Mismatches 32; Indels 5; Gaps 1;

Db 1 MDIDPKYKFGATVLLSPDPSFVSVDLDTASALREALSEPHHTRALROI 60
QY :|||||
113 MDIDPKYKFGATVLLSPDPSFVSVDLDTASALREALSEPHHTRALROI 172
Db 61 CWDDELKLIAMSSNITSEQVRTIIVNHNVDTWGKIVQSLMFHSCITFGHTQDELY 120
QY :|||||
173 CWDDELKLIAMSSNITSEQVRTIIVNHNVDTWGKIVQSLMFHSCITFGHTQDELY 232
Db 121 SPGWIRPPAPYRPPNAPILSTLPETTYVRRDRGRSPRRRSPRRRSQ 180
QY :|||||
233 SPGWIRPPAPYRPPNAPILSTLPETTYVRRG-----RSPRRRTPSPRRRSQ 287
Db 181 SOSPSANC 188
QY :|||
288 SOSRESQC 295

RESULT 14
ID CORA_HPBGS STANDARD: PRT; 217 AA.
AC P03153;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1990 (REL. 13, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS GROUND SQUIRREL HEPATITIS VIRUS (GSVH).
OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 84267998.
RA SEEGER C., GANEM D., VARMS H.E.;
RT "Nucleotide sequence of an infectious molecularly cloned genome of
RT ground squirrel hepatitis virus.";
RL J. VIROL. 51:367-375(1984).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: K02715; G325401; -.
DR PIR: A03715; NKVL5.
DR PFAM: PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 196 203
FT REPEAT 204 211
SQ SEQUENCE 217 AA; 25189 MW; B70A00B1 CRC32;

Query Match 42.4%; Score 934; DB 1; Length 217;
Best Local Similarity 68.1%; Pred. No. 3,206-146;
Matches 128; Conservative 22; Mismatches 32; Indels 6; Gaps 2;

Db 31 MDIDPYKEFGSSVQLNLFPLDFFPDNALVDYTAALYEELTGREHCSPHHTAIRQALV 90
113 MDIDPYKEFGATVLELFLPSDFPSVRLDLDITASALYREALSPHCSPHHTAIRQAIL 172
QY 91 CMDELTLITMMSNINTESEVRR-IIVDHVNNTWGLKVRQITLWHLSCITFGQHTVOEFLV 149
173 CWGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFLHISCTFGRETIVIEYLV 232
Db 150 SFVWIRTTPAPYRPPNAPILSTLPEHTVIRRGAGARSPPRRTPSPRRRSQSPRRRR 209
QY 233 SFGVWIRTTPAYRPPNAPILSTLPEHTIVRRRG-----RSPRRRTSPRRRSQSPRRRR 287
Db 210 QSPSANC 217
QY 288 QSRRESQC 295

RESULT 15
ID CORA_MHV8 STANDARD; PRT; 187 AA.
AC P06433;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 8 (MHV 8).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86062931.
RA KODAMA K., OGASAWARA N., YOSHIKAWA H., MURAKAMI S.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT evolutionary relationship between hepadnaviruses.";
RL J. VIROL. 56:978-986(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M11082; G336135; -.
DR PIR: A03714; NKVL2.
DR PFAM: PF00906; Hepatitis_core; 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 166 173
FT REPEAT 174 181
SQ SEQUENCE 187 AA; 21579 MW; A1C354F3 CRC32;

Query Match 42.3%; Score 933; DB 1; Length 187;
Best Local Similarity 66.3%; Pred. No. 4,946-146;
Matches 124; Conservative 26; Mismatches 33; Indels 4; Gaps 2;

Db 1 MDIDPYKEFGSSVQLNLFPLDFFPDNALVDYTAALYEELTGREHCSPHHTAIRQALV 60
113 MDIDPYKEFGATVLELFLPSDFPSVRLDLDITASALYREALSPHCSPHHTAIRQAIL 172
QY 61 CMDELTLITMMSNINTESEVRR-IIVDHVNNTWGLKVRQITLWHLSCITFGQHTVOEFLV 120
173 CWGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFLHISCTFGRETIVIEYLV 232
Db 121 SFVWIRTTPAPYRPPNAPILSTLPEHTVIRRGAGARSPPRRTPSPRRRSQSPRRRR 180
QY 233 SFGVWIRTTPAYRPPNAPILSTLPEHTIVRRRG-----RSPRRRTSPRRRSQSPRRRR 288
Db 181 QSPSANC 187
QY 289 QSRRESQC 295

Search completed: Thu Dec 16 13:49:59 1999
Job time : 33 secs.

(a) (b) (c) (d) (e)

Query Match 62.4%; Score 1375; DB 14; Length 212;
Best Local Similarity 99.5%; Pred. No. 2,79e-219;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171

DB 89 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 148
QY 172 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 231

DB 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 291

DB 209 ESQC 212
QY 292 ESQC 295

RESULT 2 PRELIMINARY; PRT; 183 AA.

ID 089437;
AC 089437;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (CASTA 2).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RN (1)
RC STRAIN-PATIENT CASTA-2/87;
RC LAI M.E., MAZZOLENTI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN (2)
RN (2)
RP SEQUENCE FROM N.A.
RP STRAIN-AYW;
RC KARAYIANNIS P.;
RA SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN (3)
RN (3)
RP SEQUENCE FROM N.A.
RP STRAIN-AYW;
RC KARAYIANNIS P.;
RA SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85289; G736113; -;
DR EMBL: X80925; E198085; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 62.3%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 6,48e-219;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 60
QY 113 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 172

DB 61 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 120
QY 173 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 232

DB 121 SFGWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 180
QY 233 SFGWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 292

DB 181 SOC 183
QY 293 SOC 295

RESULT 3

ID 068008 PRELIMINARY; PRT; 183 AA.
AC 068008;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (FERACUT 1).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RN (1)
RP SEQUENCE FROM N.A.
RP STRAIN-PATIENT FERACUT-1/89;
RC LAI M.E., MAZZOLENTI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85284; G736098; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 183 AA; 21102 MW; BB9AFAFB CRC32;

Query Match 62.3%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 6,48e-219;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 60
QY 113 MDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 172

DB 61 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 120
QY 173 CMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 232

DB 121 SFGWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 180
QY 233 SFGWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 292

DB 181 SOC 183
QY 293 SOC 295

RESULT 4 PRELIMINARY; PRT; 212 AA.

ID 068020;
AC 068020;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN (1)
RN (1)
RP SEQUENCE FROM N.A.
RP STRAIN-PATIENT CASTAG-1/85;
RC LAI M.E., MAZZOLENTI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85291; G736119; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 62.3%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4,25e-219;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPYKEFGATVLLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171

DB 89 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 148
QY 172 LCMGELMTLATWGVNLEDPASRDVLSYVNTNMGKFRQLMFHISCLTFGRETIVIELY 231

DB 149 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPEITVVRGRGSPRRRTSPRRRSQSPRRRSQSR 291

Db 209 ESQC 212
1111
QY 292 ESQC 295

RESULT 5
ID 069597 PRELIMINARY; PRT; 212 AA.
AC 069597;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN-SUB-TYPE AYM.
RA MEDLINE; 94079539.
RA PREISLER-ADAMS S., SCHLAYER M.J., PETERS T., HETTLER F., GEROK W.,
RA RASENACK J.;
RT "Sequence analysis of hepatitis B virus DNA in immunologically
negative infection";
RL ARCH. VIROL. 133:385-396(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYM;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYM;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X172702; G288830; -;
DR EMBL: X80925; E198084; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24336 MW; 575D20BF CRC32;

Query Match 62.3%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4.25e-219;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 88
:|||||
QY 112 SMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 171

Db 89 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 148
:|||||
QY 172 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 231

Db 149 VSEGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
:|||||
QY 232 VSEGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 291

Db 209 ESQC 212
1111
QY 292 ESQC 295

RESULT 6
ID 067876 PRELIMINARY; PRT; 212 AA.
AC 067876;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CT, HBV SUBTYPE AYM;
RA LAI M.E., MAZOLENTI A.P., BALESTRIERI A., MELIS A., PORRU A.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65258; G59436; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;

Query Match 62.3%; Score 1373; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 6.48e-219;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 88
:|||||
QY 112 SMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 171

Db 89 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 148
:|||||
QY 172 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 231

Db 149 VAFGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
:|||||
QY 232 VSEGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 291

Db 209 ESQC 212
1111
QY 292 ESQC 295

RESULT 7
ID 011884 PRELIMINARY; PRT; 212 AA.
AC 011884;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN PRECURSOR.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA RAO B.S., CASEY J.L., RINAUDO J.S., KORBA B.E.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U95551; G2182120; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;

Query Match 62.2%; Score 1372; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 9.87e-219;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 88
:|||||
QY 112 SMDIDPKKEFGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHNTALROAI 171

Db 89 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 148
:|||||
QY 172 LCMGELMTLATWGVNLEDPAASRDVLSYVNTNMGKFRQLMFIISCLTFGRETIVIEL 231

Db 149 VSEGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 208
:|||||
QY 232 VSEGVWIRTPPAYRPPNAPILISTLPETTVARRGRSPRRRTSPRRRSOSPRRRRSOSR 291

Db 209 EPOC 212
1111
QY 292 ESQC 295

RESULT 8
ID 068066 PRELIMINARY; PRT; 183 AA.
AC 068066;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (LICHERI 2).
OS CORE.
OS HEPATITIS B VIRUS.

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2'87;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -
PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 183 AA; 21102 MW; 6F38AA3B CRC32;

Query Match 62.1%; Score 1370; DB 14; Length 183;
Best Local Similarity 98.4%; Pred. No. 2.29e-218;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 60
QY 113 MDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 172
Db 61 CWDMLTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 120
QY 173 CWDMLTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 232
Db 121 SFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
QY 233 SFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 292
Db 181 SOC 183
QY 293 SOC 295

RESULT 9
ID 068068 PRELIMINARY; PRT; 212 AA.
AC 068068;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3'90;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -
PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;

Query Match 62.1%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.29e-218;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
QY 112 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 231
Db 149 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 232 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 10
ID 068025 PRELIMINARY; PRT; 212 AA.

AC 068025;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG'3;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -
PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;

Query Match 62.1%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.29e-218;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
QY 112 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 231
Db 149 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 232 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 11
ID 067872 PRELIMINARY; PRT; 212 AA.
AC 067872;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C, HBSAG POSITIVE, HBV SUBTYPE ATW;
RA LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -
PFAM: PF00906; Hepatitis_core; 1.
SO SEQUENCE 212 AA; 24348 MW; A6BD9DA1 CRC32;

Query Match 62.1%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 3.49e-218;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 88
QY 112 GMDIDPKKEGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALRQAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFQRLMFHISCLTFGREYIEYL 231
Db 149 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 232 VSFQVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSR 291
Db 209 ESOC 212

```
QY 292 ESOC 295
|||||
RESULT 12
ID 067984 PRELIMINARY; PRT; 212 AA.
AC 067984;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT VITTORINA'92;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85256; G736052; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24322 MW; 83D9780B CRC32;

Query Match 62.1%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3.49e-218;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 231
Db 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 13
ID 068032 PRELIMINARY; PRT; 212 AA.
AC 068032;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CHIGINE-2'86;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85296; G736139; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24308 MW; 69D87B53 CRC32;

Query Match 62.1%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 3.49e-218;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 231
Db 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 14
ID 068010 PRELIMINARY; PRT; 212 AA.
AC 068010;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24292 MW; 04A4D12D CRC32;

Query Match 62.0%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 5.32e-218;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 231
Db 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 15
ID 068070 PRELIMINARY; PRT; 212 AA.
AC 068070;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GIORDO'84;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24376 MW; 80F52D0F CRC32;

Query Match 62.0%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.24e-217;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 231
Db 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 14
ID 068010 PRELIMINARY; PRT; 212 AA.
AC 068010;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24292 MW; 04A4D12D CRC32;

Query Match 62.0%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 5.32e-218;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 112 SMDIDPKFEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 171
Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 148
QY 172 LCMGELMTLATWGVNLEDPASRDVSVYNTNGLFROLMHISCLTFRGTVEYL 231
Db 149 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 208
QY 232 VSGFWIRTPPAYRPPNAPILSTLPTTVVRRGRSPRRRTSPRRRSOSPRRRRSOSR 291
Db 209 ESOC 212
QY 292 ESOC 295

RESULT 15
ID 068070 PRELIMINARY; PRT; 212 AA.
AC 068070;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GIORDO'84;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210; -.
PFAM: PF00906; Hepatitis_core: 1.
SQ SEQUENCE 212 AA; 24376 MW; 80F52D0F CRC32;

Query Match 62.0%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1.24e-217;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 29 GMDIDPYKEGATVELLSFLPSDFPSPVRLDLDTASALYREALESPHCSPHHTALROAI 88
:|||||
QY 112 SMDIDPYKEGATVELLSFLPSDFPSPVRLDLDTASALYREALESPHCSPHHTALROAI 171
|||||
Db 89 LCMGELMTLATWGVNLEDPASRDVYSYNTNMGKFRQLMWFHISCLTFGRETVEYL 148
|||||
QY 172 LCMGELMTLATWGVNLEDPASRDVYSYNTNMGKFRQLMWFHISCLTFGRETVEYL 231
|||||
Db 149 VSFGVWIRTPPAYRPPNAPILTLPETTVVRRGRSPRRTPSPRRRSQSPRRRSQSR 208
|||||
QY 232 VSFGVWIRTPPAYRPPNAPILTLPETTVVRRGRSPRRTPSPRRRSQSPRRRSQSR 291
|||||
Db 209 ESQC 212
|||||
QY 292 ESQC 295
```

Search completed: Thu Dec 16 13:52:01 1999
Job time : 104 secs.

Thu Dec 16 12:41:41 1999

us3890-1-38183.pcp

Page 1

US3890-1-38183
MDIDPYKEFGATVELLSFLPSDFPVSVDLDTASALYREALSEPHCSPHHTALROAILCWEIMTLAT
WGVNLEDPASRDLYSVNEITRDGFLLOMDGFEPEHLVDFLOSMDIDPYKEFGATVELLSFLP
SDPEVSVDLDTASALYREALSEPHCSPHHTALROAILCWEIMTLATWGVNLEDPASRDLYSVN
TNMGLKFRQLMFIHSCITFCREIVIELVSGWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRT
PSPRRRSOSFRRRSOSRESOC1

 W O R L D
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch.p protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:32:26 1999; Maspar time 13.80 Seconds

Tabular output not generated. 467.015 Million cell updates/sec

Title: >US3890-1-38183
 Description: (1-303) from us3890-1-38183.pep
 Perfect Score: 2263
 Sequence: 1 MDIDYKEFGATVELLSFLP.....RRRSQSPRRRSQSNESQC 303

Scoring table: PAM 150
 Gap 11

Searched: 170751 seqs, 21266608 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39

Statistics: Mean 33.516; Variance 173.687; scale 0.193

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	60.8	194	32	W50242	Hepatitis B virus pre
2	1375	60.8	212	32	W50250	Hepatitis B virus pre
3	1375	60.8	346	5	R27473	S12/core protein.
4	1363	60.2	184	1	P80959	Hepatitis B virus sub
5	1359	60.1	193	32	W50241	Hepatitis B virus pre
6	1358	60.0	183	32	W50251	Hepatitis B virus p21
7	1353	59.8	183	20	W09044	Hepatitis B virus cor
8	1349	59.6	183	5	R05635	Hepatitis B antigen.
9	1344	59.4	397	20	W09041	Sequence of core anti
10	1338	59.1	184	5	P00004	Plasmid pHBV DN AA en
11	1337	59.1	208	8	R40808	Sequence of core anti
12	1337	59.0	196	8	R40806	Hepatitis B core / PV
13	1336	58.9	183	4	P40311	Hepatitis B core / PV
14	1334	58.9	183	4	R40805	Hepatitis B virus core
15	1334	58.9	183	4	R40805	Hepatitis B virus core
16	1334	58.9	183	13	R68868	Hepatitis B virus pol

Result	ID	Score	Query Match	Length	ID	Description	Pred. No.
17	1334	58.9	183	12	R62869	Hepatitis B virus cor	1.55e-96
18	1333	58.9	198	8	R40807	Hepatitis B core / IL	1.88e-96
19	1328	58.7	293	10	R55286	Presl full length cor	4.83e-96
20	1325	58.6	183	13	R68866	Hepatitis B virus pol	8.51e-96
21	1325	58.6	214	1	P90702	Deduced amino acid se	8.51e-96
22	1325	58.6	289	20	W09049	Plasmid pHBV DN BB en	8.51e-96
23	1324	58.5	185	6	R30861	Hepatitis B core anti	1.03e-95
24	1324	58.5	185	10	R55284	Deduced sequence of f	1.03e-95
25	1324	58.5	185	6	R30780	Native HBcAg protein.	1.03e-95
26	1324	58.5	185	6	R31025	Native HBcAg protein.	1.03e-95
27	1321	58.4	185	26	P30051	Core antigen of HBV.	1.81e-95
28	1320	58.3	183	13	R68869	Hepatitis B virus pol	2.19e-95
29	1317	58.2	185	6	R30781	Altered HBcAg protein	3.85e-95
30	1317	58.2	185	6	R31026	Altered HBcAg protein	3.85e-95
31	1314	58.1	183	13	R68867	Hepatitis B virus pol	6.79e-95
32	1314	58.0	183	13	R68870	Hepatitis B virus pol	8.20e-95
33	1311	57.9	185	6	R30784	Altered HBcAg protein	1.20e-94
34	1311	57.9	185	6	R31029	Altered HBcAg protein	1.20e-94
35	1311	57.9	185	6	R30865	Modified hepatitis B	1.20e-94
36	1309	57.8	185	6	R30862	Hepatitis B core anti	1.75e-94
37	1308	57.8	185	6	R31028	Altered HBcAg protein	2.11e-94
38	1308	57.8	185	6	R30783	Altered HBcAg protein	2.11e-94
39	1308	57.8	185	6	R30864	Modified hepatitis B	2.11e-94
40	1303	57.6	214	1	P80961	HBV core antigen enco	5.42e-94
41	1302	57.5	203	1	P82872	HBcAg/Beta-Gal fusion	6.55e-94
42	1287	56.9	198	32	W50252	Hepatitis B virus pre	1.11e-92
43	1263	55.8	183	17	R98878	Hepatitis B virus E a	1.03e-90
44	1259	55.6	183	17	R98883	Hepatitis B virus E a	2.19e-90
45	1258	55.6	183	17	R98884	Hepatitis B virus E a	2.64e-90

ALIGNMENTS

RESULT 1
 ID W50242 standard: Protein; 194 AA.

AC W50242:
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p22 polypeptide Met-p22.
 KW Viral replication; inhibitor; HBV; nucleocapsid; gene therapy;
 KM hepatocyte; liver; Met-p22.
 OS Hepatitis B virus.
 OS Synthetic.
 FH Key
 FT Protein
 FT Location/Qualifiers
 FT 2..194
 FT /Label= p22
 FT W0909649-A1.
 PD 12-MAR-1998.
 PF 03-SEP-1997; U15500.
 PR 03-SEP-1996; US-025370.
 PA (GENE) GEN HOSPIRAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 98-193325/17
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 11: Page 40: 60pp: English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 22 kDa (p22)
 CC protein with an added N-terminal Met residue. p22 is produced by
 CC elimination of the 19-amino acid leader peptide from the 25 kDa
 CC full-length HBV precore protein (see W50250). Evidence is provided
 CC that HBV replication is inhibited in the presence of high levels of
 CC HBV precore or precore-related proteins. These proteins can be
 CC incorporated into HBV nucleocapsids along with the p21 core protein
 CC (see W50251), which is the usual nucleocapsid component, and
 CC thereby render the nucleocapsids deficient in encapsidating HBV
 CC pregenomic RNA. Thus, over-expression of the precore proteins, or
 CC certain variants of them, leads to transdominant inhibition of HBV
 CC replication. Suitable inhibitory proteins include p25 (see W50250),
 CC p22, Met-p22, p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). Heterologous peptides (see W50244-49) may be
 CC inserted into the p22 and Met-p22 polypeptides. The inhibitory
 CC proteins can be produced by recombinant methods using claimed
 CC expression vectors and host cells. They can be provided exogenously

CC to the target cells for use in inhibiting HBV replication.
 CC Alternatively, a nucleic acid construct that directs overexpression
 CC of an inhibitory protein in target cells is used for the gene
 CC therapy of HBV infection.
 SQ Sequence 194 AA;

Query Match 60.8%; Score 1375; DB 32; Length 194;
 Best Local Similarity 99.5%; Pred. No. 6.72e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 11 gmdidykefgatvellsfipstfipsvrldltaasalyrealsepncsphtalrgai 70
 :|||||
 QY 120 SMDIDPYKEGATVELLSFIPSDFFPSVRDLDTASALYREALSEPHCSPHHTALROAI 179
 Db 71 lwcgelmtlatwgvnlgedpasrdlvsvyntmglkfifqllwfhlscltfgretvleyl 130
 :|||||
 QY 180 LCGELMTLATWGVNLGEDPASRDLVSVYNTMGLKFRQLMFHISCLTFCGRETVEYL 239
 Db 131 vsfgvwlrtppayrppnapilstlpettvrrrgsprrrrrsgsprrrrrsgsr 180
 :|||||
 QY 240 VSFQVWIRTPPAYRPPNAPILSTLPETTVRRRGSRPRRRTPSPRRRSQSRRRSQSR 299
 Db 191 esgc 194
 :||||
 QY 300 ESQC 303

RESULT 2

ID W50250 standard; Protein; 212 AA.
 AC W50250;
 DT 28-SEP-1998 (first entry)
 DE Hepatitis B virus precore p25 polypeptide.
 KM Hepatitis B virus; precore; p25 polypeptide; gene therapy;
 OS Hepatocyte; liver; p25 protein.
 KM Hepatitis B virus.
 OS Hepatitis B virus.
 PN W09809649-A1.
 PD 12-MAR-1998.
 PF 03-SEP-1997; 015500.
 PR 03-SEP-1996; US-025370.
 PA (GHEO) GEN HOSPITAL CORP.
 PI Melegari M, Scaglioni PP, Wands JR;
 DR WPI: 98-193325/17.
 PT DNA encoding proteins which can be incorporated with wild type
 PT nucleocapsid subunit(s) into a viral nucleocapsid - useful for
 PT inhibition of viral replication, especially hepatitis B virus
 PS Claim 15; Page 35; 60pp; English.
 CC This polypeptide comprises the hepatitis B virus (HBV) 25 kDa (p25)
 CC protein that is encoded by the full-length HBV precore gene.
 CC Evidence is provided that HBV replication is inhibited in the
 CC presence of high levels of HBV precore or precore-related proteins.
 CC These proteins can be incorporated into HBV nucleocapsids along
 CC with the p21 core protein (see W50251), which is the usual
 CC nucleocapsid component, and thereby render the nucleocapsids
 CC deficient in encapsidating HBV pregenomic RNA. Thus, over-
 CC expression of the precore proteins, or certain variants of them,
 CC leads to transdominant inhibition of HBV replication. Suitable
 CC inhibitory proteins include p25, p22 (see W50241), Met-p22 (see
 CC W50242), p18 (see W50236), Met-p18 (see W50237) and Met-p18-Het
 CC (see W50238). The inhibitory proteins can be produced by
 CC recombinant methods using claimed expression vectors and host
 CC cells. They can be provided exogenously to the target cells for
 CC use in inhibiting HBV replication. Alternatively, a nucleic acid
 CC construct that directs overexpression of an inhibitory protein in
 CC target cells is used for the gene therapy of HBV infection.
 SQ Sequence 212 AA;

Query Match 60.8%; Score 1375; DB 32; Length 212;
 Best Local Similarity 99.5%; Pred. No. 6.72e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 29 gmdidykefgatvellsfipstfipsvrldltaasalyrealsepncsphtalrgai 88
 :|||||
 QY 120 SMDIDPYKEGATVELLSFIPSDFFPSVRDLDTASALYREALSEPHCSPHHTALROAI 179

Db 89 lwcgelmtlatwgvnlgedpasrdlvsvyntmglkfifqllwfhlscltfgretvleyl 148
 :|||||
 QY 180 LCGELMTLATWGVNLGEDPASRDLVSVYNTMGLKFRQLMFHISCLTFCGRETVEYL 239
 Db 149 vsfgvwlrtppayrppnapilstlpettvrrrgsprrrrrsgsprrrrrsgsr 208
 :|||||
 QY 240 VSFQVWIRTPPAYRPPNAPILSTLPETTVRRRGSRPRRRTPSPRRRSQSRRRSQSR 299
 Db 209 esgc 212
 :||||
 QY 300 ESQC 303

RESULT 3

ID R27473 standard; Protein; 346 AA.
 AC R27473;
 DT 24-FEB-1993 (first entry)
 DE S12/core protein.
 KM Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
 KM Vaccinia virus; 13L; promoter; NVVAC; recombinant; HBV L;
 KM Large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
 KM S1, S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
 KM deletion loci; recipient loci.
 OS Synthetic.
 FS Key
 FT region 1..108
 FT region /label S1
 FT region 109..163
 FT region /label S2
 FT region 164..346
 FT region /label Core
 PN W09215672-A.
 PD 17-SEP-1992.
 PF 09-MAR-1992; U01906.
 PR 07-MAR-1991; US-666056.
 PR 11-JUN-1991; US-713967.
 PR 06-MAR-1992; US-847951.
 PA (VIRO-) VIRGENETICS CORP.
 PI Cox WI, De Talsne C, Francis J, Gettig RR, Johnson GP,
 PI Limpach KJ, Norton EK, Paolletti E, Perkins ME, Plancus SE,
 PI Riviere M, Tartaglia J, Taylor J;
 DR WPI: 92-331718/40.
 DR N-PSDB: Q29105.
 PT Vaccine comprises recombinant, attenuated pox-virus - use for
 PT vaccinating against viral infections such as rabies, hepatitis B,
 PT HIV, HSV, EBV, CMV, mumps etc.
 PS Disclosure: fig 13; 45pp; English.
 CC The sequence given is encoded by an expression cassette which
 CC comprises a portion of the hepatitis B virus (HBV) pre-S region (S12/
 CC core) which is precisely linked to the 13L promoter derived from
 CC vaccinia virus. This DNA sequence was linked to the S1 and S2
 CC sequences and this fragment was used in the construction of a NVVAC
 CC recombinant expressing the HBV gene. Other HBV genes were also used
 CC in the construction. These were HBV M protein (small pre-S antigen,
 CC spsAg) and HBV L (large pre-S antigen, lpsAg). Each of these gene
 CC sequences were inserted individually into three different sites of
 CC NVVAC separated by from each other by large regions of vaccinia DNA
 CC containing essential genes. NVVAC is a Copenhagen vaccine strain of
 CC vaccinia virus which has been modified by deletion of six non-essential
 CC regions of the genome encoding known or potential virulence factors.
 CC The deletion loci were engineered as recipient loci for the insertion
 CC of foreign genes. The spacing of the three inserted sequences ensured
 CC that any recombination that did occur would lead to disruption of the
 CC vaccinia genome and would cause unviable vaccinia virus. See also
 CC Q35501-864.
 SQ Sequence 346 AA;

Query Match 60.8%; Score 1375; DB 5; Length 346;
 Best Local Similarity 99.5%; Pred. No. 6.72e-100;
 Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 163 mndidykefgatvellsfipstfipsvrldltaasalyrealsepncsphtalrgai 222

PF 1-JAN-1989; 123526.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOJ) Biogen Inc.
 PI Murray K, Schaller HE;
 DR WPI: 90-195067/26.
 N-PSDB: 004799
 PT Recombinant DNA encoding Hepatitis B virus polypeptide antigens -
 PS used in detection of infection and in vaccine prodn.
 PS Disclosure; 4pp; English.
 CC Polypeptide fragments encoded by the DNA sequence are antigenic for
 CC HBV and may be used as a vaccine or in detection. Peptides may be
 CC cultured in a suitable bacterial host such as E.coli.
 CC Fragments of the sequence are also claimed as being antigenically
 CC useful.
 SO Sequence 183 AA;

Query Match 59.6%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 9.14e-98;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdldpykefgatvellsfipsdffpsrvdlldtaaalrydalespeshphltarail 60
 |||||
 QY 121 MDIDPYKEFGATVELLSFLPSDFPSRVDLDTASALYREALSPESHPHLTARAIL 180
 Db 61 cwgdlmclatwgvnledpsardlvsvyvtnglktfqlwfhscitfgrevleylv 120
 |||||
 QY 181 CWGELMTLATWGVNLEDPSARDLVSVYVTNGLKTFQLWFMHISCITFGREVLEYLV 240
 Db 121 sfgywiltppayrpnapllstlpctvrrrgsprrrpsrrrrsgsre 180
 |||||
 QY 241 SFGWIRTPPAYRPNAPILSTLPETVRRGRSPRRRSPRRRSQSRRRSQSRE 300
 Db 181 sqc 183
 ||||
 QY 301 SQC 303

RESULT 9

ID P00041 standard; Protein; 183 AA.
 AC P00041;
 DT 14-OCT-1992 (first entry)
 DE Sequence of core antigen.
 KM Hepatitis B virus; antigen: antibody; diagnosis; vaccine.
 OS Hepatitis B virus.
 PS EP-13828-A.
 PD 06-AUG-1980.
 PF 21-DEC-1979; 303017.
 PR 22-DEC-1978; GB-049907.
 PR 27-DEC-1978; GB-050039.
 PR 01-NOV-1979; GB-037910.
 PA (BIOJ) Biogen Inc.
 PI Murray K, Schaller HE;
 DR WPI: 80-57286C/33.
 N-PSDB: N00003.
 PT Recombinant DNA coding for polypeptide - have specificity of
 PT hepatitis B viral antigens in detection or antibody stimulation
 PS Example: Figs 3-4; 43pp; English.
 CC Human serum from a single HBsAg positive, HBeAg positive donor
 CC (serotype adym) was used to prep. a DNA-contg. pellet which was
 CC labelled with 3H or 32P as described by P. M. Kaplan et al (1973).
 CC The labelled DNA was then extracted with phenol from the resulting
 CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).
 CC It was then cloned in plasmid pBR322 which was used to transform E.
 CC coli. Micro-organisms prepd. by the processes are deposited at the
 CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg;
 CC HBV-Kpn I dc; Tetr Amps HBV+.
 SO Sequence 183 AA;

Query Match 59.6%; Score 1349; DB 5; Length 183;
 Best Local Similarity 96.7%; Pred. No. 9.14e-98;
 Matches 177; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 1 mdldpykefgatvellsfipsdffpsrvdlldtaaalrydalespeshphltarail 60
 |||||
 QY 121 MDIDPYKEFGATVELLSFLPSDFPSRVDLDTASALYREALSPESHPHLTARAIL 180
 Db 61 cwgdlmclatwgvnledpsardlvsvyvtnglktfqlwfhscitfgrevleylv 120
 |||||
 QY 181 CWGELMTLATWGVNLEDPSARDLVSVYVTNGLKTFQLWFMHISCITFGREVLEYLV 240
 Db 121 sfgywiltppayrpnapllstlpctvrrrgsprrrpsrrrrsgsre 180
 |||||
 QY 241 SFGWIRTPPAYRPNAPILSTLPETVRRGRSPRRRSPRRRSQSRRRSQSRE 300
 Db 181 sqc 183
 ||||
 QY 301 SQC 303

Query Match 59.4%; Score 1344; DB 20; Length 397;
 Best Local Similarity 100.0%; Pred. No. 2.35e-97;
 Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mdldpykefgatvellsfipsdffpsrvdlldtaaalrydalespeshphltarail 60
 |||||
 QY 121 MDIDPYKEFGATVELLSFLPSDFPSRVDLDTASALYREALSPESHPHLTARAIL 180
 Db 61 cwgdlmclatwgvnledpsardlvsvyvtnglktfqlwfhscitfgrevleylv 120
 |||||
 QY 181 CWGELMTLATWGVNLEDPSARDLVSVYVTNGLKTFQLWFMHISCITFGREVLEYLV 240
 Db 121 sfgywiltppayrpnapllstlpctvrrrgsprrrpsrrrrsgsre 179
 |||||

Query 241 SFGWIRTPPAYRPPNAPILSTLPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSR 299

RESULT 11
ID P00004 standard; Protein: 184 AA.

AC P00004;
DT 14-OCT-1992 (first entry)
DE Sequence of core antigen.
OS Hepatitis B virus; antigen; antibody; diagnosis; vaccine.

PN EP-13828-A.
PD 06-AUG-1980.
PF 21-DEC-1979: 303017.
PR 22-DEC-1978: GB-049907.
PR 27-DEC-1978: GB-050039.
PR 01-NOV-1979: GB-037910.
PA (BIOT) BIOGEN NV.
PI Murray K, Schaller HE.
DR WPI: 80-57268C/33.
DR N-PSDB; N00002.
PT Recombinant DNA coding for polypeptide - have specificity of
PT Hepatitis B viral antigens in detection or antibody stimulation
PS Clam 13; Page 40; 43pp; English.

CC (serotype adym) was used to prep. a DNA-contg. pellet which was
CC Human serum from a single HBSAg positive, HBeAg positive donor
CC labeled with 3H or 32P as described by P. M. Kaplan et al (1973).

CC The labeled DNA was then extracted with phenol from the resulting
CC pellet using the procedure of L. I. Lutwick and W. S. Robinson (1977).

CC It was then cloned in plasmid pBR322 which was used to transform E.
CC coli. Micro-organisms prep. by the processes are deposited at the
CC NCIB as pBR322-HBV-G-L, e.g. E. coli HB101/pBR322-Pst I dg:
CC HBV-kpn I dc: Tetr Amps HBV+.

SO Sequence 184 AA:

Query Match 59.1%; Score 1338; DB 5; Length 184;

Best Local Similarity 96.2%; Pred. No. 7.30e-97;

Matches 177; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

DB 1 mdidpykefagavellslfspdffpsvrcdlldtaaalrydalespeshcphhtalrgai 60

QY 121 MDIDPYKEFGATVELLSFLPSDFPSV-RDLDTASALYREALSPHCHPHHTALRQAI 179

DB 61 lcwgdmlatwvgnlndpasrdlvsvyvmnglklrqlwfhiscldfgrvleyl 120

QY 180 LCWGEMLTATWGVNLEDPASRDLVSVYVNMGLKFRQLWFHISCLTFGREVLEYL 239

DB 121 vsfgwlrtppayrppnapilslpettvrrrgisprrrpsrrrrsqsprrrrsqr 180

QY 240 VSEFGWIRTPPAYRPPNAPILSTLPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSR 299

DB 181 esgc 184

QY 300 ESGC 303

RESULT 12
ID R40808 standard; Protein: 208 AA.

AC R40808;
DT 16-FEB-1994 (first entry)
DE Hepatitis B core / PV-1 / IL-1 Fusion.
KM Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.
OS Synthetic.

FT Key Location/Qualifiers

FT region 8..19 /label= PV-1

FT region 20..25 /label= IL-1

FT protein /label= HBC

FT J05192170-A.
PN 03-AUG-1993.
PD 24-SEP-1991: 243800.

PR 24-SEP-1991: JP-243800.

PA (NISM) NISSHIN OIL MILLS LTD.

DR WPI: 93-277479/35.

DR N-PSDB: 047738.

PT Recombinant plasmid for high immunogenicity virus - contains

PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia

PT virus and exotic genes

PS Disclosure; Fig 9; 12pp; Japanese.

CC The hepatitis B core gene is recombined with PV-1 DNA

CC and IL-1 beta to form a plasmid (047738) which is then used to produce

CC a vaccine. NB: Sequence is difficult to read in the original

CC specification.

SO Sequence 208 AA:

Query Match 59.1%; Score 1337; DB 8; Length 208;

Best Local Similarity 94.6%; Pred. No. 8.82e-97;

Matches 175; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

DB 24 lsmdidpykefagavellslfspdffpsvrcdlldtaaalrydalespeshcphhtalrgai 83

QY 119 LSMDIDPYKEFGATVELLSFLPSDFPSV-RDLDTASALYREALSPHCHPHHTALRQAI 178

DB 84 lcwgdmlatwvgnlndpasrdlvsvyvmnglklrqlwfhiscldfgrvleyl 143

QY 179 LCWGEMLTATWGVNLEDPASRDLVSVYVNMGLKFRQLWFHISCLTFGREVLEYL 238

DB 144 lvsfgwlrtppayrppnapilslpettvrrrgisprrrpsrrrrsqsprrrrsqr 203

QY 239 LVSEFGWIRTPPAYRPPNAPILSTLPETTVVARRGRSPRRRTSPRRRSOSPRRRRSOSR 298

DB 204 resgc 208

QY 299 RESGC 303

RESULT 13

ID R40806 standard; Protein: 196 AA.

AC R40806;
DT 16-FEB-1994 (first entry)

DE Hepatitis B core / PV-1 fusion.

KM Hepatitis B; core; HBC; PV-1; IL-1 beta; interleukin; plasmid.

OS Synthetic.

FT Key Location/Qualifiers

FT region 7..13 /label= PV-1

FT protein /label= HBC

PN J05192170-A.

PD 03-AUG-1993.

PF 24-SEP-1991: 243800.

PR 24-SEP-1991: JP-243800.

PA (NISM) NISSHIN OIL MILLS LTD.

DR WPI: 93-277479/35.

DR N-PSDB: 047736.

PT Recombinant plasmid for high immunogenicity virus - contains

PT recombinant haemagglutinin gene, hepatitis B core gene of vaccinia

PT virus and exotic genes

PS Disclosure; Fig 7; 12pp; Japanese.

CC The hepatitis B core gene is recombined with PV-1 DNA

CC and IL-1 beta to form a plasmid (047738) which is then used to produce

CC a vaccine. NB: Sequence is difficult to read in the original

CC specification.

SO Sequence 196 AA:

Query Match 59.0%; Score 1336; DB 8; Length 196;

Best Local Similarity 95.1%; Pred. No. 1.07e-96;

Matches 175; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 13 amdidpykefagavellslfspdffpsvrcdlldtaaalrydalespeshcphhtalrgai 72

QY 120 AMDIDPYKEFGATVELLSFLPSDFPSV-RDLDTASALYREALSPHCHPHHTALRQAI 179

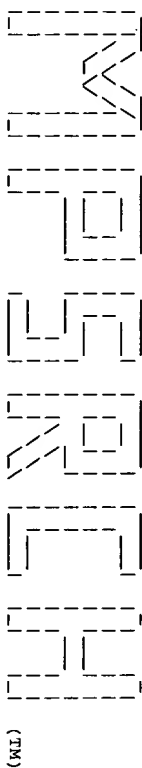
DB 73 lcwgdmlatwvgnlndpasrdlvsvyvmnglklrqlwfhiscldfgrvleyl 132

PT virus and exotic genes

```

search completed: 11th Dec 10 13:33:27 1999
Job time : 181 secs.

```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Dec 16 13:35:44 1999; Maspar time 14.81 Seconds
820.017 Million cell updates/sec
Tabular output not generated.

Title: >US3890-1-38183
Description: (1-303) from us3890-1-38183.pep
Perfect Score: 2263
Sequence: 1 MDIDYKEGATVELLSFLP.....RRRRSQSPRRRSQSNESQC 303

Scoring table: PAM 150
Gap 11

Searched: 122810 seqs, 40068593 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: plr60
1:plr1 2:plr2 3:plr3 4:plr4

Statistics: Mean 47.447; Variance 117.780; scale 0.403

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	1375	60.8	212	1	NKVLAH	e antigen precursor / 1.94e-202	
2	1373	60.7	183	2	S53207	core antigen - hepati / 4.22e-202	
3	1374	60.7	212	2	S53211	e antigen precursor / 2.66e-202	
4	1374	60.7	212	2	S53204	e antigen precursor / 2.66e-202	
5	1373	60.7	212	2	S20750	e antigen precursor / 4.22e-202	
6	1370	60.5	183	2	S53270	core antigen - hepati / 1.35e-201	
7	1370	60.5	212	2	S53216	e antigen precursor / 1.35e-201	
8	1370	60.5	212	2	S53272	e antigen precursor / 1.35e-201	
9	1369	60.5	212	2	S53163	e antigen precursor / 2.01e-201	
10	1369	60.5	212	2	S20746	e antigen precursor / 2.01e-201	
11	1369	60.5	212	2	S53225	e antigen precursor / 2.01e-201	
12	1368	60.5	212	2	S53198	e antigen precursor / 2.96e-201	
13	1366	60.4	212	2	S53223	e antigen precursor / 6.45e-201	
14	1366	60.4	212	2	S53274	e antigen precursor / 6.45e-201	
15	1364	60.3	183	2	S53169	core antigen - hepati / 1.41e-200	
16	1363	60.2	183	2	S53247	core antigen - hepati / 2.08e-200	
17	1363	60.2	212	2	S53202	e antigen precursor / 2.08e-200	
18	1363	60.2	212	2	S53159	e antigen precursor / 2.08e-200	
19	1363	60.2	212	2	S53200	e antigen precursor / 3.07e-200	
20	1362	60.2	212	2	S53251	e antigen precursor / 3.07e-200	
21	1360	60.1	212	2	S53204	e antigen precursor / 6.68e-200	
22	1360	60.1	212	2	S53227	e antigen precursor / 6.68e-200	
23	1359	60.1	212	2	S53253	e antigen precursor / 9.87e-200	

24	1358	60.0	211	1	NKVLAI	e antigen precursor / 1.46e-199
25	1357	60.0	212	2	S53281	e antigen precursor / 2.15e-199
26	1355	59.9	212	1	NKVLBH	e antigen precursor / 4.69e-199
27	1355	59.9	212	2	S53242	e antigen precursor / 4.69e-199
28	1351	59.7	183	2	S53129	core antigen - hepati / 2.23e-198
29	1351	59.7	212	2	S53240	core antigen - hepati / 2.23e-198
30	1349	59.6	183	1	NKVLAI	e antigen precursor / 4.85e-198
31	1349	59.6	183	2	S53181	core antigen - hepati / 4.85e-198
32	1348	59.6	183	2	S53232	core antigen - hepati / 7.16e-198
33	1348	59.6	183	2	S53260	core antigen - hepati / 7.16e-198
34	1345	59.4	183	2	S53152	core antigen - hepati / 2.30e-197
35	1344	59.4	212	2	S53229	e antigen precursor / 3.40e-197
36	1343	59.3	183	2	S53214	core antigen - hepati / 5.02e-197
37	1342	59.3	183	2	S53140	core antigen - hepati / 7.40e-197
38	1341	59.3	183	2	S53267	core antigen - hepati / 1.09e-196
39	1341	59.3	183	2	S53189	core antigen - hepati / 1.09e-196
40	1341	59.3	212	2	S53238	e antigen precursor / 1.09e-196
41	1339	59.2	183	2	S53137	core antigen - hepati / 2.38e-196
42	1340	59.2	212	2	S53257	e antigen precursor / 1.61e-196
43	1339	59.2	212	2	S52651	e antigen precursor / 2.38e-196
44	1338	59.1	212	2	S53236	e antigen precursor / 3.51e-196
45	1337	59.1	212	1	NKVLJ2	e antigen precursor / 5.19e-196

ALIGNMENTS

RESULT 1
ENTRY NKVLAH #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw4, isolate hb321 and others)
ALTERNATE_NAMES HBE antigen precursor / HBC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal name hepatitis B virus; HBV
#subtype ayw4, isolate hb321; isolate patient Ferracuti '83;
isolate patient Castag '83; isolate patient Sanna '84;
isolate patient Licheri '1'85; isolate patient Flore '1'86;
isolate patient Licheri '83
18-Dec-1981 #sequence_revision 08-Nov-1996 #text_change
10-Oct-1997
DATE
ACCESSIONS S47405; S53191; S53209; S53234; S53264; S53249; S53262;
S53277; A03711
REFERENCE S47404
#authors Plucienniczak, A.
#submission Submitted to the EMBL Data Library, August 1994
#description Molecular cloning and sequencing of two complete genomes of
polish isolates of human hepatitis B virus.
#accession S47405
#molecule_type DNA
#residues 1-212 #label PLU
#cross-references EMBL:235716; NID:9527435; PID:9527437
#experimental_source subtype ayw4, isolate hb321
REFERENCE S53112
#authors Lai, M.E.; Marzotenti, A.P.; Porru, A.; Balestrieri, A.
#submission Submitted to the EMBL Data Library, March 1995
#accession S53191
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:X85283; NID:g736088; PID:g736090
#experimental_source isolate patient Ferracuti '83
#accession S53209
#molecule_type DNA
#residues 1-212 #label LA2
#cross-references EMBL:X85290; NID:g736114; PID:g736116
#experimental_source isolate patient Castag '83
#accession S53234
#molecule_type DNA
#residues 1-212 #label LA3
#cross-references EMBL:X85300; NID:g736150; PID:g736152
#experimental_source isolate patient Sanna '84
#accession S53264
#molecule_type DNA
#residues 1-212 #label LA4
#cross-references EMBL:X85313; NID:g736194; PID:g736196

	#experimental_source isolate patient Lichert-1'85
#accession	S53249
##molecule_type	DNA
##residues	1-212 ##label LA5
##cross-references	EMBL:X85306; NID:g736172; PID:g736174
##experimental_source	isolate patient Flore-1'86
#accession	S53262
##molecule_type	DNA
##residues	1-212 ##label LA6
##cross-references	EMBL:X85312; NID:g736191; PID:g736193
##experimental_source	isolate patient Licheri'83
#accession	S53277
##molecule_type	DNA
##residues	30-212 ##label LA7
##cross-references	EMBL:X85317; NID:g736211; PID:g736214
##experimental_source	patient Giorio-2'86
#note	due to a stop codon between the alternative initiators the e antigen precursor cannot be processed
REFERENCE	A93214
##authors	Gallibert, F.; Mandart, E.; Fitoussi, F.; Tiollais, P.; Charnay, P.
##journal	Nature (1979) 281:646-650
##title	Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.
##cross-references	MUID:81012091
#accession	A03711
##molecule_type	DNA
##residues	1-212 ##label GAL
##cross-references	GB:J02203; NID:g329640; PID:g329642
##experimental_source	subclone ayw
GENETICS	
##gene	C
CLASSIFICATION	##superfamily hepatitis B virus core antigen
KEYWORDS	alternative initiators; core protein
FEATURE	
1-29	domain signal sequence #status predicted #label SIG\
30-212	product core antigen #status predicted #label CAG\
30-178	product e antigen #status predicted #label EAG\
179-212	domain carboxyl-terminal prepeptide #link EAG #status predicted #label ECP
SUMMARY	#length 212 #molecular_weight 24350 #checksum 782
Query Match	60.8%; Score 1375; DB 1; Length 212;
Best Local Similarity	99.5%; Pred. No. 1.gae-202;
Matches 185; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
Db	29 GMDIDPYKEFGATVELLSFLPSDEFFPSYRDLDDASALYREALSPHRCSPHHNTALROAI 88
Oy	: : : : : : : : : :
	120 SMDIDPYKEFGATVELLSFLPSDEFFPSYRDLDDASALYREALSPHRCSPHHNTALROAI 179
Db	89 LCWGLMLTLATWGVGNLEDPASRDILVSYVTNNGLKFRQLMWHICLTFGRTVEIYL 148
Oy	: : : : : : : : :
	180 LCWGLMLTLATWGVGNLEDPASRDILVSYVTNNGLKFRQLMWHICLTFGRTVEIYL 239
Ddb	149 VSFQWITTPAYRPNPAPILSTLPETTVARRQRSPRRPTSPRRRSQSPPRRRSQSR 208
Oy	: : : : : : : :
	240 VSFGWITTPAYRPNPAPILSTLPETTVARRQRSPRRPTSPRRRSQSPPRRRSQSR 299
Dd	209 ESOC 212
Oy	
	300 ESOC 303
RESULT	2
ENTRY	S53207 #type complete
TITLE	core antigen - hepatitis B virus (isolate patient Castaa-2'87)
ALTERNATE_NAMES	HBC antigen
CONTAINS	core antigen
ORGANISM	#formal name hepatitis B virus, HBV
#variety	isolate patient Castaa-2'87
DATE	08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

```

ACCESSIONS      S53207
REFERENCE        S53112
#authors        Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission     Submitted to the EMBL Data Library, March 1995
#accession      S53207
##molecule-type DNA
##residues      1-183 ##label LAI
##cross-references EMBL:X85289; NID:g736110; PID:g736113
##experimental_source isolate patient Castaa-2/87
#note           due to a stop codon between the alternative initiators
                the e antigen precursor cannot be produced

GENETICS
#gene           C
CLASSIFICATION  #superfamily hepatitis B virus core antigen
KEYWORDS        core protein
SUMMARY         #length 183 #molecular-weight 21102 #checksum 2165

Query Match      60.7%; Score 1373; DB 2; Length 183;
Best Local Similarity 99.5%; Pred. No. 4,22e-202;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDYKEFGATVLLSFLSPDEFPSYRDLDTASALYRDALSPHCSPHNTALRAIL 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 121 MDIDYKEFGATVLLSFLSPDEFPSYRDLDTASALYRDALSPHCSPHNTALRAIL 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 61 CWGELMTLATVGVNLDSPASRDLYSVYVNTNNGLFROLWPHISCLTFGRRTVEYLY 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 181 CWGELMTLATVGVNLDSPASRDLYSVYVNTNNGLFROLWPHISCLTFGRRTVEYLY 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 121 SFGWIRTPPARPPNRPILSTLPETVTVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 241 SFGWIRTPPARPPNRPILSTLPETVTVRRGRSPRRRTSPRRRSQSPRRRSQSR 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 181 SQC 183
| | |
Oy 301 SQC 303

RESULT          3
ENTRY           S53211 #type complete
TITLE           e antigen precursor / core antigen - hepatitis B virus
                (isolate patient Castag-1/85 and others)
ALTERNATE_NAMES HBC antigen; HBe antigen precursor / HBC antigen; pre-C/C
                antigen
CONTAINS        core antigen; e antigen
ORGANISM        #format_name hepatitis B virus, HBV
                #variety isolate patient Castag-1/85; isolate patient Ferracuti-1/89
                08-Jul-1995 #sequence-revision 03-Aug-1995 #text_change
                08-Sep-1997
DATE            S53211; S53197
ACCESSIONS      S53112
REFERENCE        S53112
#authors        Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission     Submitted to the EMBL Data Library, March 1995
#accession      S53211
##molecule-type DNA
##residues      1-212 ##label LAI
##cross-references EMBL:X85291; NID:g736117; PID:g736119
##experimental_source isolate patient Castag-1/85
#accession      S53197
##molecule-type DNA
##residues      30-212 ##label LAN
##cross-references EMBL:X85284; NID:g736095; PID:g736098
##experimental_source isolate patient Ferracuti-1/89
#note           due to a stop codon between the alternative initiators
                the e antigen precursor cannot be produced

GENETICS
#gene           C
CLASSIFICATION  #superfamily hepatitis B virus core antigen
KEYWORDS        alternative initiators; core protein
FEATURE         #domain signal sequence #status predicted #label SIG\
                30-212 #product core antigen #status predicted #label CAG\
                30-178 #product e antigen #status predicted #label EAG\

```


179-212 #domain carboxyl-terminal propeptide #link EAG #status
predicted #label ECP
SUMMARY #length 212 #molecular-weight 24336 #checksum 746

Query Match
Best Local Similarity 98.9%; Pred. No. 2,866-202;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 88
QY 120 SMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 179

Db 89 LCMGDLTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 148
QY 180 LCMGELMTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 239

Db 149 VSEGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 208
QY 240 VSEGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 299

Db 209 ESOC 212
QY 300 ESOC 303

RESULT 4
ENTRY S32204 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1000)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus; HBV
#subtype ayw, patient C1000
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997

ACCESSIONS
REFERENCE S32202
S32204
PRESTIER-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.
#submitted to the EMBL Data Library, March 1993
#description Identification and sequence analysis of hepatitis B virus DNA in immunological negative infection.
#accession S32204
#molecule_type DNA
#residues 1-212 #label PRE
#cross-references EMBL:X72702; NID:g288927; PID:g288930
#experimental_source subtype ayw, patient C1000
#note due to a stop codon between the alternative initiators the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
FEATURE 1-29
1-29
30-212 #domain signal sequence #status predicted #label SIG
30-178 #product core antigen #status predicted #label CAG
179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24336 #checksum 770

Query Match
Best Local Similarity 98.9%; Pred. No. 2,866-202;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 88
QY 120 SMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 179

Db 89 LCMGDLTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 148
QY 180 LCMGELMTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 239

Db 149 VSEGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 208
QY 240 VSEGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 299

Db 209 ESOC 212
QY 300 ESOC 303

RESULT 5
ENTRY S20750 #type complete
TITLE e antigen precursor / core antigen - hepatitis B virus
(subtype ayw, patient C1)
ALTERNATE_NAMES HBe antigen precursor / HBeC antigen; pre-C/C antigen
CONTAINS core antigen; e antigen
ORGANISM #formal_name hepatitis B virus; HBV
#subtype ayw, patient C1
DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997

ACCESSIONS
REFERENCE S20750
S20745
PORRU, A.
#submitted to the EMBL data library, March 1992
#description Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic liver disease.
#accession S20750
#molecule_type DNA
#residues 1-212 #label LAI
#cross-references EMBL:x65258; NID:g59434; PID:g59436
#experimental_source subtype ayw, patient C1

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS alternative initiators; core protein
FEATURE 1-29
30-212 #domain signal sequence #status predicted #label SIG
30-178 #product core antigen #status predicted #label CAG
179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 147

Query Match
Best Local Similarity 98.9%; Pred. No. 4,226-202;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 88
QY 120 SMDIDPKKEGATVELLSFLPSDFPSVRLDLPASALYREALSPHCSPHHTALRQAI 179

Db 89 LCMGDLTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 148
QY 180 LCMGELMTLATWGVNEDPASRDLYSVYNTNGLKROLMLHISCLTGRETVIEL 239

Db 149 VAFGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 208
QY 240 VSEGVWIRTPAPYPPNAPILSTLPETTVARRGRSPRRRTPSPRRRSQSR 299

Db 209 ESOC 212
QY 300 ESOC 303

RESULT 6
ENTRY S53270 #type complete
TITLE e antigen - hepatitis B virus (isolate patient Licheri-2/87)
ALTERNATE_NAMES HBe antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus; HBV
#isolate patient Licheri-2/87
DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change

08-Sep-1997

ACCESSIONS S53270

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53270

##molecule_type DNA

##residues 1-183 ##label LAI

##cross-references EMBL:X85314; NID:9736201; PID:9736204

##experimental_source isolate patient Licheri-2'87

##note due to a stop codon between the alternative initiators the e antigen precursor cannot be produced

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS core protein

SUMMARY #length 183 #molecular-weight 21102 #checksum 2199

Query Match 60.5%; Score 1370; DB 2; Length 183;

Best Local Similarity 98.9%; Pred. No. 1.36e-201;

Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 MDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 60

QY 121 MDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 180

Db 61 CWDMLTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 120

QY 181 CWDMLTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 240

Db 121 SFGVWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 180

QY 241 SFGVWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 300

Db 181 SOC 183

QY 301 SOC 303

RESULT 7

ENTRY S53216 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES (isolate patient Castag'3)

CONTAINS core antigen precursor / Hbc antigen; pre-C/C antigen

ORGANISM core antigen; e antigen

##formal_name hepatitis B virus, HBV

#variety isolate patient Castag'3

DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53216

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53216

##molecule_type DNA

##residues 1-212 ##label LAI

##cross-references EMBL:X85293; NID:9736124; PID:9736126

##experimental_source isolate patient Castag'3

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS alternative initiators; core protein

FEATURE

1-29

30-212 #domain signal sequence #status predicted #label SIG\

30-178 #product e antigen #status predicted #label CAG\

179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24363 #checksum 752

Query Match 60.5%; Score 1370; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 1.36e-201;

Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 88

QY 120 SMDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 179

Db 89 LCMGELMTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 148

QY 180 LCMGELMTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 239

Db 149 VSEGWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 208

QY 240 VSEGWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 299

Db 209 ESOC 212

QY 300 ESOC 303

RESULT 8

ENTRY S53272 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES (isolate patient Licheri-3'90)

CONTAINS core antigen precursor / Hbc antigen; pre-C/C antigen

ORGANISM core antigen; e antigen

##formal_name hepatitis B virus, HBV

#variety isolate patient Licheri-3'90

DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53272

REFERENCE S53112

#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.

#submission submitted to the EMBL Data Library, March 1995

#accession S53272

##molecule_type DNA

##residues 1-212 ##label LAI

##cross-references EMBL:X85315; NID:9736205; PID:9736207

##experimental_source isolate patient Licheri-3'90

GENETICS

#gene C

CLASSIFICATION #superfamily hepatitis B virus core antigen

KEYWORDS alternative initiators; core protein

FEATURE

1-29

30-212 #domain signal sequence #status predicted #label SIG\

30-178 #product e antigen #status predicted #label CAG\

179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24366 #checksum 446

Query Match 60.5%; Score 1370; DB 2; Length 212;

Best Local Similarity 98.9%; Pred. No. 1.36e-201;

Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 88

QY 120 SMDIDPKKEGATVELLSFLPSDFPSVDLDTASALYREALSPHCSPHHTALROAI 179

Db 89 LCMGELMTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 148

QY 180 LCMGELMTLATWGVNLDPASRDVSVYNTNMGKLFROLMFHISCLTFGRVIEYL 239

Db 149 VSEGWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 208

QY 240 VSEGWIRPPAYRPNPAILSTLPETTVARRGRSPRRRSPRRRSOSR 299

Db 209 ESOC 212

QY 300 ESOC 303

RESULT 9

ENTRY S53163 #type complete

TITLE e antigen precursor / core antigen - hepatitis B virus

ALTERNATE_NAMES (isolate patient Vitorina/92)
 CONTAINS HBe antigen precursor / Hbc antigen: pre-C/C antigen
 #formal_name hepatitis B virus, HBV
 #variety isolate patient Vitorina/92
 DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53163
 REFERENCE S53112
 #authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
 #submission submitted to the EMBL Data Library, March 1995
 #accession S53163
 #molecule_type DNA
 #residues 1-212 #label LAI
 #cross-references EMBL:X85256; NID:g736050; PID:g736052
 #experimental_source isolate patient Vitorina/92

GENETICS C
 #gene #superfamily hepatitis B virus core antigen
 #key_words alternative initiators; core protein
 FEATURE 1-29
 30-212 #domain signal sequence #status predicted #label SIG\
 30-178 #product core antigen #status predicted #label CAG\
 179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24322 #checksum 842

Query Match 60.5%; Score 1369; DB 2; Length 212;
 Best Local Similarity 98.9%; Pred. No. 2,01e-201;
 Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 :|||||
 QY 120 SMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179
 :|||||

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 148
 :|||||
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 239
 :|||||

Db 149 VSGFWITPPAYRPPNPILSTLPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 208
 :|||||
 QY 240 VSGFWITPPAYRPPNPILSTLPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 299
 :|||||

Db 209 ESQC 212
 :|||||
 QY 300 ESQC 303

RESULT 10
 ENTRY S20746 #type complete
 TITLE e antigen precursor / core antigen - hepatitis B virus
 (subtype ayw, patient C)
 ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
 CONTAINS core antigen; e antigen
 ORGANISM #formal_name hepatitis B virus, HBV
 #variety subtype ayw, patient C
 DATE 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997

ACCESSIONS S20746
 REFERENCE S20745
 #authors Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.
 #submission submitted to the EMBL Data Library, March 1992
 #description Sequence analysis of HBV genomes isolated from patients with HBsAg negative chronic liver disease.

#accession S20746
 #molecule_type DNA
 #residues 1-212 #label LAI
 #cross-references EMBL:X65257; NID:g59429; PID:g59431
 #experimental_source subtype ayw, patient C

GENETICS C
 #gene

CLASSIFICATION #superfamily hepatitis B virus core antigen
 KEYWORDS alternative initiators; core protein
 FEATURE 1-29
 30-212 #domain signal sequence #status predicted #label SIG\
 30-178 #product core antigen #status predicted #label CAG\
 179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24348 #checksum 506

Query Match 60.5%; Score 1369; DB 2; Length 212;
 Best Local Similarity 98.4%; Pred. No. 2,01e-201;
 Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 :|||||
 QY 120 SMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179
 :|||||

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 148
 :|||||
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 239
 :|||||

Db 149 VSGFWITPPAYRPPNPILSTLPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 208
 :|||||
 QY 240 VSGFWITPPAYRPPNPILSTLPETTVRRRGRSPRRRTSPRRRSQSPRRRSQSR 299
 :|||||

Db 209 ESQC 212
 :|||||
 QY 300 ESQC 303

RESULT 11
 ENTRY S53225 #type complete
 TITLE e antigen precursor / core antigen - hepatitis B virus
 (isolate patient Chigline-2/86)
 ALTERNATE_NAMES HBe antigen precursor / Hbc antigen: pre-C/C antigen
 CONTAINS core antigen; e antigen
 ORGANISM #formal_name hepatitis B virus, HBV
 #variety isolate patient Chigline-2/86
 DATE 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 08-Sep-1997

ACCESSIONS S53225
 REFERENCE S53112
 #authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
 #submission submitted to the EMBL Data Library, March 1995
 #accession S53225
 #molecule_type DNA
 #residues 1-212 #label LAI
 #cross-references EMBL:X85256; NID:g736137; PID:g736139
 #experimental_source isolate patient Chigline-2/86

GENETICS C
 #gene #superfamily hepatitis B virus core antigen
 #key_words alternative initiators; core protein
 FEATURE 1-29
 30-212 #domain signal sequence #status predicted #label SIG\
 30-178 #product core antigen #status predicted #label CAG\
 179-212 #domain carboxyl-terminal propeptide #link EAG #status predicted #label ECP

SUMMARY #length 212 #molecular-weight 24308 #checksum 655

Query Match 60.5%; Score 1369; DB 2; Length 212;
 Best Local Similarity 98.4%; Pred. No. 2,01e-201;
 Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 :|||||
 QY 120 SMDIDPKFEGATVLLFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179
 :|||||

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 148
 :|||||
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKRFQRLMFHISCLTFCGRTVIEYL 239
 :|||||

Best Local Similarity 98.9%; Pred. No. 6,45e-201;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPKYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 88
QY 120 SMDIDPKYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEL 239
Db 149 VSGVWIRTPPAYRPPNAPILITLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILITLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 15
ENTRY S53169 #type complete
TITLE core antigen - hepatitis B virus (isolate patient Muresu'89)
ALTERNATE_NAMES HBC antigen
CONTAINS core antigen
ORGANISM #formal_name hepatitis B virus, HBV
#isolate_patient Muresu'89
#isolate_patient Muresu'89
#sequence_revision 03-Aug-1995 #text_change 08-Sep-1997
DATE S53169
ACCESSIONS S53112
REFERENCE S53112
#authors Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
#submission submitted to the EMBL Data Library, March 1995
#accession S53169
#molecule_type DNA
##residues 1-183 #label LAI
##cross_references EMBL:X85275; NID:g736057; PID:g736060
##experimental_source isolate patient Muresu'89
#note due to a stop codon between the alternative initiators the e antigen precursor cannot be produced

GENETICS
#gene C
CLASSIFICATION #superfamily hepatitis B virus core antigen
KEYWORDS core protein
SUMMARY #length 183 #molecular-weight 21088 #checksum 1849

Query Match 60.3%; Score 1364; DB 2; Length 183;
Best Local Similarity 98.4%; Pred. No. 1.41e-200;
Matches 180; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 1 MDIDPKYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 60
QY 121 MDIDPKYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTALROAI 180
Db 61 CMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEL 120
QY 181 CMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTFGRETIVIEL 240
Db 121 SFGVWIRTPPAYRPPNAPILITLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
QY 241 SFGVWIRTPPAYRPPNAPILITLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 300
Db 181 SOC 183
QY 301 SOC 303

Search completed: Thu Dec 16 13:37:00 1999
Job time : 76 secs.

(TM)

11cm

100

3

E-14

E-14

Matches 183: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 1 MDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTALQAIL 60
  |||
Qy 121 MDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTALQAIL 180
Db 61 CMGELMTLATWGVNLEDPASRDLYVSYNTNMGLKFRQLMFWHISCLTFGRETVEYLY 120
  |||
Qy 181 CMGELMTLATWGVNLEDPASRDLYVSYNTNMGLKFRQLMFWHISCLTFGRETVEYLY 240
Db 121 SEGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 180
  |||
Qy 241 SEGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 300
Db 181 SOC 183
  |||
Qy 301 SOC 303

RESULT 2
ID CORA_HPBVA STANDARD: PRT: 211 AA.
AC P24023:
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (STRAIN ALPHA1).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90266476.
RA TONG S., LI J., VITVITSKI L., TREPO C.;
RT "Active hepatitis B virus replication in the presence of anti-HBe is
  associated with viral variants containing an inactive pre-C region.";
RL VIROLOGY 176:596-603(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M32138; G495034; ALT_SEQ.
DR PIR: A34773; NKVLA1.
DR PFAM: PF00906; Hepatitis_core; 1.
DR CORE PROTEIN; REPEAT.
FT REPEAT 183 190
FT REPEAT 198 206
SQ SEQUENCE 211 AA: 24208 MW: 0924CESE CRC32:

```

Query Match 60.0%; Score 1358; DB 1; Length 211;
 Best Local Similarity 95.2%; Pred. No. 7.55e-231;
 Matches 179: Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

Db 24 LGWLMDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTAL 83
  |||
Qy 116 LGWLMDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTAL 175
Db 84 ROAILCWGLDITLSTWGVNLEDPTSRLVSYNTNMGLKFRQLMFWHISCLTFGRETVEY 143
  |||
Qy 176 ROAILCWGLDITLSTWGVNLEDPTSRLVSYNTNMGLKFRQLMFWHISCLTFGRETVEY 245
Db 144 IEVLVSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 203
  |||
Qy 236 IEVLVSGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 295
Db 204 TOSRESOC 211
  |||
Qy 296 TOSRESOC 303

```

RESULT 3
 ID CORA_HPBVZ STANDARD: PRT: 183 AA.
 AC P03147;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
 DE CORE ANTIGEN.
 GN C.
 OS HEPATITIS B VIRUS (SUBTYPE ADW).
 OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 81012115.
 RA PASER M., GOTO T., GILBERT W., ZINK B., SCHALLER H., MCKAY P.,
 RA LEADBETTER G., MURRAY K.;
 RT "Hepatitis B virus genes and their expression in E. coli.";
 RL NATURE 282:575-579(1979).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J02202; G329638; -
 DR EMBL: A08967; G411874; -
 DR PIR: B93217; NKVLA2.
 DR PFAM: PF00906; Hepatitis_core; 1.
 DR CORE PROTEIN; REPEAT.
 FT REPEAT 162 169
 FT REPEAT 170 177
 SQ SEQUENCE 183 AA: 21042 MW: 6EA6B674 CRC32:

Query Match 59.6%; Score 1349; DB 1; Length 183;
 Best Local Similarity 96.7%; Pred. No. 4.36e-229;
 Matches 177: Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

Db 1 MDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTALQAIL 60
  |||
Qy 121 MDIDPKKFGATVELLSFLPSDFPSVRLDLDYASALYREALSPHCSPHHTALQAIL 180
Db 61 CMGELMTLATWGVNLEDPASRDLYVSYNTNMGLKFRQLMFWHISCLTFGRETVEYLY 120
  |||
Qy 181 CMGELMTLATWGVNLEDPASRDLYVSYNTNMGLKFRQLMFWHISCLTFGRETVEYLY 240
Db 121 SEGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 180
  |||
Qy 241 SEGVWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSOSRRSRE 300
Db 181 SOC 183
  |||
Qy 301 SOC 303

RESULT 4
ID CORA_HPBVO STANDARD: PRT: 183 AA.
AC P17392:
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN OKINAWA/PODW282).  

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.  

RN [1]  

RP SEQUENCE FROM N.A.  

RX MEDLINE: 89010694.  

RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEMIGNO R.I., IMAI M.,  

RA MIYAKAWA Y., MAYUMI M.;  

RT "Typing hepatitis B virus by homology in nucleotide sequence:  

  comparison of surface antigen subtypes.";
```



```

RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.",
RL J. GEN. VIROL. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V00587; -. NOT_ANNOTATED_CDS.
DR EMBL: D00331; -. NOT_ANNOTATED_CDS.
DR PIR: A93480; NKVLAS.
DR PIR: B93460; NKVLAA.
DR PIR: C28925; NKVLJ3.
DR PIR: PF00906; Hepatitis_core.1.
DR CORE_PROTEIN: REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
FT SEQUENCE 183 AA; 21095 MW; D2E185B9 CRC32;
Query Match 58.9%; Score 1334; DB 1; Length 183;
Best Local Similarity 95.6%; Pred.No.3,75e-226;
Matches 175; Conservative 5; Mismatches 25; Indels 0; Gaps 0;
Db 1 MDIDPYKEFGAVSVELLSFLPSDFPSIRDLDTASALYREALSEPHSCPHHTALRAIL 60
QY 121 MDIDPYKEFGAVTVELLSFLPSDFPSVADLDTASALYREALSEPHSCPHHTALRAIL 180
Db 61 CWGELMNATATVGSLEDPASRELVSIVYNNMGKITQQLMFHISCLTFGEITYLEYIV 120
QY 181 CWGELMTATATVGVLEDPASRDVLVSIVNNMGKITQQLMFHISCLTFGEITYLEYIV 240
Db 121 SEGWIPTPPAIPPPNADIIISLPTETVARRGRSPRRRTSPRRRSOSPPRRRSOSRE 180
QY 241 SFGWIRTPPAIPPPNADIIISLPTETVARRGRSPRRRTSPRRRSOSPPRRRSOSRE 300
Db 181 SOC 183
QY 301 SOC 303
RESULT 6
ID CORA_HPBVJ STANDARD; PRT: 183 AA.
AC P17391;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
OS HEPATITIS B VIRUS (SUBTYPE ADW / STRAIN JAPAN/PJDM233).
NC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
CC [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89010694.
RA OKAMOTO H., TSUDA F., SAKUGAWA H., SASTROSEWIGNO R.I., IMAI M.,
RT MIYAKAWA Y., MAYUMI M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.",
RL J. GEN. VIROL. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00329; -. NOT_ANNOTATED_CDS.
DR PIR: A28925; NKVLJ1.

```

DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 162 169
FT REPEAT 170 177
SQ SEQUENCE 183 AA: 21224 MW: F6B348B6 CRC32:

Query Match 58.4%; Score 1322; DB 1; Length 183;
Best Local Similarity 96.2%; Pred. No. 8,34e-224;
Matches 176; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 60
121 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 180
Db 61 CMCGLMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 120
181 CMCGLMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 240
Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 180
241 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 300
Qy 181 SOC 183
111
Qy 301 SOC 303

RESULT 7
ID CORA_HPBVT STANDARD: PRT; 185 AA.
AC P03149;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 83168919.

RA ONO Y., ONDA H., SASADA R., IGARASHI K., SUGINO Y., NISHIOKA K.;
RT "The complete nucleotide sequences of the cloned hepatitis B virus
DNA: subtype adr and adv.";
RL NUCLEIC ACIDS RES. 11:1747-1757(1983).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----

CC EMBL: V00866; -; NOT_ANNOTATED_CDS.
DR PIR: C93460; NKVLN6.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 164 171
FT REPEAT 172 179
SQ SEQUENCE 185 AA: 21394 MW: 791E0381 CRC32:

Query Match 58.4%; Score 1321; DB 1; Length 185;
Best Local Similarity 96.2%; Pred. No. 1,31e-223;
Matches 178; Conservative 4; Mismatches 1; Indels 2; Gaps 1;

Db 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 60
121 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 180
Db 61 CMCGLMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 120
181 CMCGLMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 240
Qy 181 SOC 183
111
Qy 301 SOC 303

Db 121 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 180
241 SFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 298
Qy 181 RESOC 185
111
Qy 299 RESOC 303

RESULT 8
ID CORA_HPBVT STANDARD: PRT; 195 AA.
AC P29178;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW2 VARIANT SF).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 90169850.
RA BHAT R.A., ULRICH P.P., VYAS G.N.;
RT "Molecular characterization of a new variant of hepatitis B virus in
RT a persistently infected homosexual man.";
RL HEPATOLOGY 11:271-276(1990).
DR PIR: A37182; NKVLH3.
DR PFAM: PF00906; Hepatitis_core; 1.
KW CORE PROTEIN; REPEAT.
FT REPEAT 174 181
FT REPEAT 182 189
SQ SEQUENCE 195 AA: 22461 MW: AE3DB5F3 CRC32:

Query Match 58.1%; Score 1314; DB 1; Length 195;
Best Local Similarity 93.5%; Pred. No. 3,06e-222;
Matches 173; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Db 11 FGLIDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 70
119 LMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALQAIL 178
Db 71 IICWELMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 130
179 IICWELMTLATWGNLDEDPASRDLYVNTNGLKIRQLMWHISCLTFGRRTVEIY 238
Qy 131 IISFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 190
239 IISFGWIRTPPAYRPPNAPILSTLPETTVARRGRSPRRRTSPRRRSQSPRRRSQSR 298
Qy 191 RESOC 195
111
Qy 299 RESOC 303

RESULT 9
ID CORA_HPBVT STANDARD: PRT; 212 AA.
AC Q05495;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS HEPATITIS B VIRUS (SUBTYPE ADW4 / STRAIN BRAZIL / ISOLATE W4B).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93346970.
RA NAUMANN H., SCHAEFER S., YOSHIDA C.F.T., GASPAR A.M.C., REPP R.,
RA GERLICH W.H.;
RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT that expresses HBV surface antigen subtype adv4.";
RL J. GEN. VIROL. 74:1627-1632(1993).
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X69798; G59423; -

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN: REPEAT.

FT DOMAIN 178 204 ARG-RICH.

FT DOMAIN 184 203 3 X 5 AA REPEATS OF S-P-R-R-R.

FT REPEAT 184 188 1.

FT REPEAT 191 195 2.

FT REPEAT 199 203 3.

SEQUENCE 212 AA; 24234 MM; BCCDF263 CRC32;

Query Match 58.0%; Score 1313; DB 1; Length 212;

Best Local Similarity 92.9%; Pred. No. 4,79e-222;

Matches 171; Conservative 10; Mismatches 3; Indels 0; Gaps 0;

DB 29 GMDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 88

OY 120 SMDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 179

DB 89 LCMGELMTLASVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 148

OY 180 LCMGELMTLATVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 239

DB 149 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSP 208

OY 240 VSEGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSP 299

DB 209 ASOC 212

OY 300 ESOC 303

RESULT 10
ID CORA_HPBVL STANDARD; PRT; 183 AA.

AC P12901; 01-OCT-1989 (REL. 12, CREATED)

DT 01-OCT-1989 (REL. 12, LAST SEQUENCE UPDATE)

DT 01-APR-1990 (REL. 14, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

C. HEPATITIS B VIRUS (STRAIN LSH / CHIMPANZEE ISOLATE).

OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

OC [1]

RN SEQUENCE FROM N.A.

RX MEDLINE: 88258473.

RA VAUDIN M., WOJSTENHOLME A.J., TSIOUAVE K.N., ZUCKERMAN A.J.,

RA HARRISON T.J.;

RT "The complete nucleotide sequence of the genome of a hepatitis B

virus isolated from a naturally infected chimpanzee.";

RL J. GEN. VIROL. 69:1383-1389(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D00220; D100603; -

DR PIR: A28885; NKVLCF.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN: REPEAT.

FT REPEAT 162 169

FT REPEAT 170 177

SEQUENCE 183 AA; 20999 MM; 1AF57C9 CRC32;

Query Match 57.8%; Score 1308; DB 1; Length 183;
Best Local Similarity 94.5%; Pred. No. 4,55e-221;
Matches 173; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

DB 1 MDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 60

OY 121 MDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 180

DB 61 CWMGELMTLASVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 120

OY 181 CWMGELMTLATVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 240

DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPA 180

OY 241 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSPRE 300

DB 181 SOC 183

OY 301 SOC 303

RESULT 11
ID CORA_HPBV2 STANDARD; PRT; 185 AA.

AC P03148; 21-JUL-1986 (REL. 01, CREATED)

DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

C. HEPATITIS B VIRUS (SUBTYPE ADW2).

OS VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

OC [1]

RN SEQUENCE FROM N.A.

RA VALENZUELA P., QUIROGA M., ZAIDIYAR J., GRAY P., RUTTER W.J.;

RL (IN) FIELD B.N., JAEINISCH R., FOX C.F. (EDS.);

RL ANIMAL VIRUS GENETICS, PP. 57-70, ACADEMIC PRESS, NEW YORK (1980).

DR PIR: A94409; NKVLA3.

DR PFAM: PF00906; Hepatitis_core; 1.

KW CORE PROTEIN: REPEAT.

FT REPEAT 164 171

FT REPEAT 172 179

SEQUENCE 185 AA; E2EA3360 CRC32;

Query Match 57.7%; Score 1305; DB 1; Length 185;

Best Local Similarity 95.7%; Pred. No. 1,76e-220;

Matches 177; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

DB 1 MDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 60

OY 121 MDIDPKFEGATVELLSFLPSDFPSVVDLDTASALYREALSPHCSPHHTALROAI 180

DB 61 CWMGELMTLATVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 120

OY 181 CWMGELMTLATVGNLLEDPASRDVLYVNTNMGKIKQLMFHISCLTFGRTVLEYL 240

DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRRTSPRRRSQSPRRRSQSP 180

OY 241 SFGVWIRTPPAYRPPNAPILSTLPETTVVRR--GRSPRRRTSPRRRSQSPRRRSQSP 298

DB 181 RESOC 185

OY 299 RESOC 303

RESULT 12
ID CORA_HPBV9 STANDARD; PRT; 214 AA.

AC P17099; 01-AUG-1990 (REL. 15, CREATED)

DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)

DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN.

C.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02715; G325401; -.
DR PIR: A03715; NKVLS.
DR PFAM: PF00906; Hepatitis_core. 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 196 203
FT REPEAT 204 211
SQ SEQUENCE 217 AA; 25189 MW; B70A00B1 CRC32;

Query Match 41.3%; Score 934; DB 1; Length 217;
Best Local Similarity 68.1%; Pred. No. 1,92e-148;
Matches 188; Conservative 22; Mismatches 32; Indels 6; Gaps 2;

Db 31 MDIDPYKFGSSYQVLNPLDFFPPDLNALVDATAALYEELTGREHCSPHHTAIRQALV 90
QY 121 MDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTAIRQAIL 180
Db 91 CWELITLITMSENTEVEVR-IIVDHVNNTWGLKVRQTLMFHLSCLTFCQHTVQEFV 149
QY 181 CWGELMTLATWGVNLDPASRDVSVYNTNMGLKFRQLMFHLSCLTFCGRETVEYLY 240
Db 150 SFGVWIRTPAPRPPNAPILSTLPEHTVIRRGSGARASPRRRTPSPRRRSOSPRRRR 209
QY 241 SFGVWIRTPAPRPPNAPILSTLPEHTVIRRGSGARASPRRRTPSPRRRSOSPRRRR 295
Db 210 QSPASNC 217
QY 296 QSPASNC 303

ID CORA_MWV8 STANDARD; PRT; 187 AA.
AC P06433;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-JAN-1988 (REL. 06, LAST SEQUENCE UPDATE)
DT 01-OCT-1989 (REL. 12, LAST ANNOTATION UPDATE)
DE CORE ANTIGEN.
GN C.
OS WOODCHUCK HEPATITIS VIRUS 8 (WHV 8).
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE; 86062931.
RA KODAMA K., OGASAWARA N., YOSHIKAWA H., MURAKAMI S.;
RT "Nucleotide sequence of a cloned woodchuck hepatitis virus genome:
RT evolutionary relationship between hepadnaviruses.";
RL J. VIROL. 56:978-986(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1082; G336135; -.
DR PIR: A03714; NKVIC2.
DR PFAM: PF00906; Hepatitis_core. 1.
KM CORE PROTEIN; REPEAT.
FT REPEAT 166 173
FT REPEAT 174 181
SQ SEQUENCE 187 AA; 21579 MW; A1C354F3 CRC32;

Query Match 41.2%; Score 933; DB 1; Length 187;
Best Local Similarity 66.3%; Pred. No. 3.00e-148;
Matches 124; Conservative 26; Mismatches 33; Indels 4; Gaps 2;

Db 1 MDIDPYKFGSSYQVLNPLDFFPPDLNALVDATAALYEELTGREHCSPHHTAIRQALV 60
QY 121 MDIDPYKFGATVELLSFLPSDFPSVRDLDTASALYREALSEPHCSPHHTAIRQAIL 180
Db 61 CWELITLITMSENTEVEVR-IIVDHVNNTWGLKVRQTLMFHLSCLTFCQHTVQEFV 120
QY 181 CWGELMTLATWGVNLDPASRDVSVYNTNMGLKFRQLMFHLSCLTFCGRETVEYLY 240
Db 121 SFGVWIRTPAPRPPNAPILSTLPEHTVIRRGSGARASPRRRTPSPRRRSOSPRRRRS 180
QY 241 SFGVWIRTPAPRPPNAPILSTLPEHTVIRRGSGARASPRRRTPSPRRRSOSPRRRRS 296
Db 181 QSPASNC 187
QY 297 QSPASNC 303

Search completed: Thu Dec 16 13:37:51 1999
Job time : 34 secs.

 NWSETH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_p protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Dec 16 13:38:08 1999; MasPar time 21.36 Seconds

Tabular output not generated. 774.258 Million cell updates/sec

Title: >US3890-1-38183

Description: (1-303) from us3890-1-38183.pep

Sequence: 1 MDIDPKKFGATVELLFLP.....RRRSQSPRRRSQSRQESQC 303

Scoring table: PAM 150
Gap 11

Searched: 179066 segs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

sptrembl9
 1:sp:archaea 2:sp:bacteria 3:sp:fungi 4:sp:human
 5:sp:invertebrate 6:sp:mammal 7:sp:mhc 8:sp:organelle
 9:sp:phage 10:sp:plant 11:sp:rodent 12:sp:unclassified
 13:sp:vertebrate 14:sp:virus

Statistics: Mean 46.887; Variance 106.690; scale 0.439

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1375	60.8	212 14	089656	PRE-C/CORE.	3.10e-221
2	1373	60.7	183 14	089437	X, PREC AND C GENES (C	7.27e-221
3	1373	60.7	183 14	068008	X, PREC AND C GENES (F	7.27e-221
4	1374	60.7	212 14	068020	PRE-C/CORE.	4.75e-221
5	1374	60.7	212 14	089597	HBCAG.	4.75e-221
6	1373	60.7	212 14	067876	PRE-C/C ORF.	7.27e-221
7	1372	60.6	212 14	011884	CORE ANTIGEN PRECURSOR	1.11e-220
8	1370	60.5	212 14	068066	X, PREC AND C GENES (L	2.60e-220
9	1370	60.5	212 14	068068	PRE-C/CORE.	2.60e-220
10	1370	60.5	212 14	068025	PRE-C/C ORF.	3.98e-220
11	1369	60.5	212 14	067984	PRE-C/CORE.	3.98e-220
12	1369	60.5	212 14	068032	PRE-C/CORE.	3.98e-220
13	1369	60.5	212 14	068010	PRE-C/CORE.	3.98e-220
14	1368	60.5	212 14	068070	PRE-C/CORE.	6.09e-220
15	1366	60.4	212 14	068030	PRE-C/CORE.	1.43e-219
16	1366	60.4	212 14	068030	PRE-C/CORE.	1.43e-219
17	1364	60.3	183 14	067989	X, PREC AND C GENES (M	3.34e-219
18	1364	60.2	183 14	068048	X, PREC AND C GENES (F	5.10e-219
19	1363	60.2	212 14	068014	PRE-C/CORE.	5.10e-219
20	1363	60.2	212 14	068012	PRE-C/CORE.	5.10e-219

Result ID	Score	Query Match	Length	ID	Description	Pred. No.
21	1363	60.2	212 14	067980	PRE-C/CORE.	5.10e-219
22	1362	60.2	212 14	068051	PRE-C/CORE.	7.81e-219
23	1360	60.1	212 14	089446	PRE-C/CORE.	1.83e-218
24	1360	60.1	212 14	068016	PRE-C/CORE.	1.83e-218
25	1360	60.1	212 14	096845	COMPLETE GENOME.	1.83e-218
26	1359	60.1	212 14	068053	PRE-C/CORE.	2.79e-218
27	1357	60.0	212 14	068075	PRE-C/CORE.	6.54e-218
28	1357	60.0	212 14	068077	PRE-C/CORE.	6.54e-218
29	1355	59.9	212 14	068045	PRE-C/CORE.	5.48e-217
30	1352	59.7	183 14	089531	CORE PROTEIN.	5.48e-217
31	1351	59.7	183 14	067946	C ANTIGEN.	8.38e-217
32	1351	59.7	212 14	089719	PRE-C/CORE.	8.38e-217
33	1349	59.6	183 14	067997	X, PREC AND C GENES (B	1.96e-216
34	1348	59.6	183 14	068037	X, PREC AND C GENES (D	3.00e-216
35	1348	59.6	183 14	068060	X, PREC AND C GENES (M	4.58e-216
36	1347	59.5	212 14	067912	PRE-C/CORE PROTEIN.	1.07e-215
37	1345	59.4	183 14	067973	X, PREC AND C GENES (T	1.07e-215
38	1344	59.4	212 14	092918	PRECORE PROTEIN PRECUR	1.64e-215
39	1344	59.4	212 14	068035	PRE-C/CORE.	1.64e-215
40	1343	59.3	183 14	068023	X, PREC AND C GENES (C	2.51e-215
41	1342	59.3	183 14	067964	X, PREC AND C GENES (F	3.84e-215
42	1341	59.3	183 14	068064	X, PREC AND C GENES (L	5.87e-215
43	1341	59.3	183 14	068003	X, PREC AND C GENES (S	5.87e-215
44	1341	59.3	212 14	068042	PRE-C/CORE.	5.87e-215
45	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215

ALIGNMENTS

Result ID	Score	Query Match	Length	ID	Description	Pred. No.
1	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
2	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
3	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
4	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
5	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
6	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
7	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
8	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
9	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
10	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
11	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
12	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
13	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
14	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
15	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
16	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
17	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
18	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
19	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
20	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
21	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
22	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
23	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
24	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
25	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
26	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
27	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
28	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
29	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
30	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
31	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
32	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
33	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
34	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
35	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
36	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
37	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
38	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
39	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
40	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
41	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
42	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
43	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
44	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215
45	1341	59.3	212 14	081115	HBCAG, HBCAG.	5.87e-215

Query Match 60.8%; Score 1375; DB 14; Length 212;
Best Local Similarity 99.5%; Pred. No. 3.10e-221;
Matches 183; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 88
:|||||
QY 120 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 179
:|||||
DB 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
:|||||
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 239
:|||||
DB 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 208
:|||||
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 299
:|||||
DB 209 ESOC 212
:|||||
QY 300 ESOC 303

RESULT 2
ID 089437 PRELIMINARY; PRT; 183 AA.
AC 089437:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (CASTAA 2).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RC SEQUENCE FROM N.A.
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AYW;
RA KARAYIANNIS P.;
RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85289; G736113; -;
DR EMBL: X80925; E198085; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; 0C504D47 CRC32;

Query Match 60.7%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 7.27e-221;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 60
:|||||
QY 121 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 180
:|||||
DB 61 CMEGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 120
:|||||
QY 181 CMEGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 240
:|||||
DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 180
:|||||
QY 241 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 300
:|||||
DB 181 SOC 183
:|||||
QY 301 SOC 303

RESULT 3

ID 068008 PRELIMINARY; PRT; 183 AA.
AC 068008:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE X, PREC AND C GENES (FERRACUT 1).
GN CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RC SEQUENCE FROM N.A.
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85284; G736098; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 183 AA; 21102 MW; BB9AFB3 CRC32;

Query Match 60.7%; Score 1373; DB 14; Length 183;
Best Local Similarity 99.5%; Pred. No. 7.27e-221;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 60
:|||||
QY 121 MDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 180
:|||||
DB 61 CMEGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 120
:|||||
QY 181 CMEGLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 240
:|||||
DB 121 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 180
:|||||
QY 241 SFGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 300
:|||||
DB 181 SOC 183
:|||||
QY 301 SOC 303

RESULT 4
ID 068020 PRELIMINARY; PRT; 212 AA.
AC 068020:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG-1'85;
RA LAI M.E., MAZZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85281; G736119; -;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24336 MW; E0E4C243 CRC32;

Query Match 60.7%; Score 1374; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 4.75e-221;
Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 29 GMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 88
:|||||
QY 120 SMDIDPKKEGATVELLSFLPSDFPSVRDLDTASALYREALESPCHSPHHTALROAI 179
:|||||
DB 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
:|||||
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 239
:|||||
DB 149 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 208
:|||||
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVVRRGRSPRRTPSPRRRSOSPRRRRSOSR 299
:|||||

Db 209 ESOC 212
 QY 300 ESOC 303

RESULT 5
 ID 089597 PRELIMINARY: PRT: 212 AA.

DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE HBAG.
 OS HEPATITIS B VIRUS.
 OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RP SEQUENCE FROM N.A.
 RC STRAIN-SUB-TYPE AYW;
 RX MEDLINE: 94079539.

RA PREISLER-ADAMS S., SCHLAYER M.J., PETERS T., HETTLER F., GEROK W.,
 RA RASENACK J.;
 RT "Sequence analysis of hepatitis B virus DNA in immunologically
 negative infection."

RL ARCH. VIROL. 133:385-396(1993).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN-AYW;
 RA KARAYIANNIS P.;
 RL SUBMITTED (DEC-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AYW;

RA KARAYIANNIS P.;
 RL SUBMITTED (SEP-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: X72702; G288930; -;
 DR EMBL: X80925; E198084; -;
 DR PFAM: PF00906; Hepatitis_core: 1.

DR PFAM: PF00906; Hepatitis_core: 1.
 SQ SEQUENCE 212 AA; 24336 MW; 575D20BF CRC32;

Query Match 60.7%; Score 1374; DB 14; Length 212;
 Best Local Similarity 98.9%; Pred. No. 4,75e-221;
 Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 QY 120 SMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 148
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 239

Db 149 VSGFWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 208
 QY 240 VSGFWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 299

Db 209 ESOC 212
 QY 300 ESOC 303

RESULT 6
 ID 067876 PRELIMINARY: PRT: 212 AA.

DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE PRE C/C ORF.
 OS HEPATITIS B VIRUS.
 OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PATIENT CT, HBV SUBTYPE AYW;
 LAI M.E., MAZZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;

RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: X65258; G59436; -;
 DR PFAM: PF00906; Hepatitis_core: 1.
 SQ SEQUENCE 212 AA; 24348 MW; 6ED741BE CRC32;

Query Match 60.7%; Score 1373; DB 14; Length 212;
 Best Local Similarity 98.9%; Pred. No. 7,27e-221;
 Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 QY 120 SMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 148
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 239

Db 149 VAFGWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 208
 QY 240 VSGFWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 299

Db 209 ESOC 212
 QY 300 ESOC 303

RESULT 7
 ID 011884 PRELIMINARY: PRT: 212 AA.

DT 01-JUL-1997 (TREMREL. 04, CREATED)
 DT 01-JUL-1997 (TREMREL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE CORE ANTIGEN PRECURSOR.
 OS HEPATITIS B VIRUS.

OC VIRUSES: RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.

RA RAO B.S., CASEY J.L., RINAUDO J.S., KORBA B.E.;
 RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL: U95551; G2182120; -;
 DR PFAM: PF00906; Hepatitis_core: 1.

DR PFAM: PF00906; Hepatitis_core: 1.
 SQ SEQUENCE 212 AA; 24360 MW; 73AC92DB CRC32;

Query Match 60.6%; Score 1372; DB 14; Length 212;
 Best Local Similarity 98.9%; Pred. No. 1,11e-220;
 Matches 182; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 88
 QY 120 SMDIDPKFEGATVELLSFLPSDFPSVRLDLDTSALYREALSPHCSPHHTALROAI 179

Db 89 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 148
 QY 180 LCMGELMTLATWGVNLEDPASRDVSVYNTNMGKFRQLMFIISCLTGTGRETIVIEL 239

Db 149 VSGFWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 208
 QY 240 VSGFWIRTPPAYRPPNAPILSTLPETTVVRRGSPRRRTSPRRRSOSPRRRRSOSR 299

Db 209 ESOC 212
 QY 300 ESOC 303

RESULT 8
 ID 068066 PRELIMINARY: PRT: 183 AA.

DT 01-NOV-1996 (TREMREL. 01, CREATED)
 DT 01-NOV-1996 (TREMREL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMREL. 08, LAST ANNOTATION UPDATE)

DE X, PREC AND C GENES (LICHERI 2).
 GN CORE.
 OS HEPATITIS B VIRUS.

OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-2/87;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85314; G736204; -
PRAM: PF00906; Hepatitis-core; 1
SQ SEQUENCE 183 AA; 21102 MW; 6F38AA3B CRC32;
Query Match 60.5%; Score 1370; DB 14; Length 183;
Best Local Similarity 98.4%; Pred. No. 2.60e-220;
Matches 180; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Db 1 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 60
QY 121 MDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 180
Db 61 CWGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 120
QY 181 CWGDLMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 240
Db 121 SFGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 180
QY 241 SFGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 300
Db 181 SQC 183
QY 301 SQC 303
RESULT 9 PRELIMINARY; PRT: 212 AA.
ID 068068 AC 068068;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN HEPATITIS B VIRUS.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT LICHERI-3/90;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85315; G736207; -
PRAM: PF00906; Hepatitis-core; 1
SQ SEQUENCE 212 AA; 24366 MW; AB227A9B CRC32;
Query Match 60.5%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.60e-220;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 29 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 88
QY 120 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 239
Db 149 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303
RESULT 10 PRELIMINARY; PRT: 212 AA.
ID 068025

AC 068025;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CASTAG/3;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85293; G736126; -
PRAM: PF00906; Hepatitis-core; 1
SQ SEQUENCE 212 AA; 24363 MW; A6A254DF CRC32;
Query Match 60.5%; Score 1370; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 2.60e-220;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 29 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 88
QY 120 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 239
Db 149 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303
RESULT 11 PRELIMINARY; PRT: 212 AA.
ID 067872 AC 067872;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE C/C ORF.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT C, HBSAG POSITIVE, HBV SUBTYPE AYW;
RA LAI M.E., MAZOLENI A.P., BALESTRIERI A., MELIS A., PORRU A.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X65257; G59431; -
PRAM: PF00906; Hepatitis-core; 1
SQ SEQUENCE 212 AA; 24348 MW; A6BD9DA1 CRC32;
Query Match 60.5%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 3.98e-220;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 29 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 88
QY 120 GMDIDPYKEGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALRQAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKLFROLMFHISCLTFGRETVIEYL 239
Db 149 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGWIRTPPAYRPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212

```
QY 300 ESOC 303

RESULT 12
ID 067984 PRELIMINARY: PRT: 212 AA.
AC 067984:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT VITTORINA'92;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85256; G736052;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24322 MM; 83D9780B CRC32;

Query Match 60.5%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 3,986-220;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 120 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 239
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 13
ID 068032 PRELIMINARY: PRT: 212 AA.
AC 068032:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT CHIGINE-2'86;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85296; G736139;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24308 MM; 69D87B53 CRC32;

Query Match 60.5%; Score 1369; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 3,986-220;
Matches 181; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 120 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 239
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 14
ID 068010 PRELIMINARY: PRT: 212 AA.
AC 068010:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24292 MM; 04A4D12D CRC32;

Query Match 60.5%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 6,096-220;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 120 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 239
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 15
ID 068070 PRELIMINARY: PRT: 212 AA.
AC 068070:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GIORDO'84;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24376 MM; 80F52D0F CRC32;

Query Match 60.4%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1,436-219;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 239
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 14
ID 068010 PRELIMINARY: PRT: 212 AA.
AC 068010:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT FERRACUTI-2'90;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85285; G736100;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24292 MM; 04A4D12D CRC32;

Query Match 60.5%; Score 1368; DB 14; Length 212;
Best Local Similarity 98.4%; Pred. No. 6,096-220;
Matches 181; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 29 GMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 88
QY 120 SMDIDPYKEFGATVELLSFLPSDFPSVRDLDTASALYREALSPHCSPHHTALROAI 179
Db 89 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 148
QY 180 LCMGELMTLATWGVNLEDPASRDLYSVYNTNMGKFRQLMFHISCLTGRETVIEYL 239
Db 149 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 208
QY 240 VSGVWIRTPPAYRPPNAPILSTLPETTVRRGRSPRRRTSPRRRSQSPRRRSQSR 299
Db 209 ESOC 212
QY 300 ESOC 303

RESULT 15
ID 068070 PRELIMINARY: PRT: 212 AA.
AC 068070:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PRE-C/CORE.
GN PRE-C/CORE.
OS HEPATITIS B VIRUS.
OC VIRUSES; RETROID VIRUSES; HEPADNAVIRIDAE; ORTHOHEPADNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PATIENT GIORDO'84;
RA LAI M.E., MAZOLENI A.P., PORRU A., BALESTRIERI A.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X85316; G736210;
DR PFAM: PF00906; Hepatitis_core; 1.
SQ SEQUENCE 212 AA; 24376 MM; 80F52D0F CRC32;

Query Match 60.4%; Score 1366; DB 14; Length 212;
Best Local Similarity 98.9%; Pred. No. 1,436-219;
Matches 182; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```

Db      29 GMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTFASALYREALESPENHSPHHTALROAI 88
      : |||||
QY      120 SMDIDPYKEFGATVELLSFLPSDFPSVRLDLDTFASALYREALESPENHSPHHTALROAI 179
      : |||||
Db      89 LCWGEIMTLATWGVNLDPASRDIVSYVNTNMGKFRQLWFIHISCLTFGRETVEYL 148
      : |||||
QY      180 LCWGEIMTLATWGVNLDPASRDIVSYVNTNMGKFRQLWFIHISCLTFGRETVEYL 239
      : |||||
Db      149 VSGVWIRTPPAYRPPNAPILTLBETVVRGRSPRRRTPSPRRRSQSPRRRSQSR 208
      : |||||
QY      240 VSGVWIRTPPAYRPPNAPILTLBETVVRGRSPRRRTPSPRRRSQSPRRRSQSR 299
      : |||||
Db      209 ESQC 212
      : |||||
QY      300 ESQC 303

```

Search completed: Thu Dec 16 13:40:01 1999
 Job time : 113 secs.



Creation date: 12-04-2003
Indexing Officer: JFUNSTEN - JAMES FUNSTEN
Team: OIPEBackFileIndexing
Dossier: 09387340

Legal Date: 07-31-2001

No.	Doccode	Number of pages
1	FOR	218

Total number of pages: 218

Remarks:

Order of re-scan issued on

